TESTS FOR CORRELATION AND PARTIAL CORRELATION BASED ON KENDALL'S TAU

BY

BASIL SAMARA

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL
OF THE UNIVERSITY OF FLORIDA IN
PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR
THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

1985

To my wife Rose, our daughter Myriam,

and to both of our families

for their love, encouragement and support

ACKNOWLEDGMENTS

I would like to thank Professor Ronald Randles for proposing the problem and for his generous guidance, patience and support throughout the research period. I would also like to thank Professors Malay Ghosh and Bill Parr for their patient assistance and helpful remarks during this time. I also thank Professors Robert Long, Dennis Wackerly and P.V. Rao for serving on my committee, and everyone at the Department of Statistics who helped in one way or another in making my stay pleasant and fruitful.

TABLE OF CONTENTS

Abstract of Dissertation Presented to the Graduate School of the University of Florida in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

TESTS FOR CORRELATION AND PARTIAL CORRELATION BASED ON KENDALL'S TAU

BY

BASIL SAMARA

August, 1985

Chairman: Dr. Ronald Randles
Major Department: Statistics

This study investigates properties of tests based on Pearson's correlation coefficient and Kendall's tau, the two most widely used measures of correlation. The main problem of interest is the partial correlation problem where the variables Y and Z are related through another variable, the covariate X. In this work each of Y and Z is related to X through the models

$$Y = \alpha_1 + \beta_1 X + E$$

and

$$Z = \alpha_2 + \beta_2 X + E' .$$

The hypotheses of interest are

1) H_0: E and E' are independent,

and

2) H_0: $\tau = 0$,

where τ is Kendall's correlation coefficient between E and E'.

For the first hypothesis, Kendall's tau calculated on the residuals from estimates of the above models, is proposed. The properties of this statistic and its asymptotic efficiency relative to the Pearson partial correlation coefficient are discussed. Also, the simulated distribution of this statistic under the null hypothesis of independence is tabulated.

The null hypothesis $\tau = 0$ is first investigated under the ordinary correlation setting between Y and Z, i.e., in the absence of the covariate term X. Here, a test is proposed based on the usual Kendall's tau but standardized by a variance estimator which has better properties than the estimators discussed in the literature. The simulated null distribution of this statistic is also given.

For the partial correlation formulation using a null hypothesis $\tau = 0$, a statistic is proposed which is similar to one studied for the ordinary correlation problem except that it is applied to the residuals from the fitted model. The simulated null distributions of this statistic generated from residuals obtained by the least squares model estimates and by least absolute regression, respectively, are also tabulated.

Results of a Monte Carlo study investigating the performances of the above statistics indicate that

(i) for hypotheses of independence, tests based on Pearson's statistics are highly robust in both the ordinary correlation, and the partial correlation settings, and that

(ii) in both settings, the tests based on our proposed modifications of Kendall's tau perform the best overall for the hypothesis that $\tau = 0$.

CHAPTER ONE
INTRODUCTION

Let (X,Y,Z) denote a random variable from some absolutely continuous trivariate distribution with distribution function F, and consider testing the null hypothesis that Y and Z are independent. If this hypothesis is rejected, one tends to believe that the variables Y and Z are dependent. However, it is possible that this "dependence" between Y and Z is due to the effect of another variable X to which both Y and Z are related in some fashion. If, for example, Y is a variable measuring mathematical ability and Z is a variable measuring musical ability, then a significant correlation between Y and Z is perhaps due to the correlation of each of Y and Z with another variable X which measures intelligence. If one suspects that such a relationship exists, then a more appropriate test may be what is commonly known as the test for partial correlation, where the null hypothesis is given by

$$H_0: \text{ Y and Z are independent conditional on X being held constant.} \tag{1.1}$$

That is to say, one "partials out" the effect of the variable X while testing the independence between Y and Z.

Although, in general, almost any relational structure between Y and X and between Z and X is possible; we use linear models as the underlying structure relating these variables. That is, we let

$$Y = \alpha_1 + \beta_1 X + E$$

and (1.2)

$$Z = \alpha_2 + \beta_2 X + E' ,$$

where the regression parameters α_1, α_2, β_1 and β_2 are unknown constants, and the random variable X is independent of both variables E and E'. Our choice of the linear structure was dictated by the fact that the normal theory procedures discussed in our work assume such a structure. For example, the use of Pearson's partial correlation coefficient (to be discussed later) is inappropriate unless both Y and Z have linear regressions on X (see, for example, Quade, 1974, p. 376 and Korn, 1984, p. 62). Under the linear models given in (1.2), the hypothesis of (1.1) is equivalent to

$$H_0: \text{ E and E' are independent .} \tag{1.3}$$

The most popular test of partial correlation is that based on Pearson's partial correlation coefficient commonly denoted by $R_{YZ.X}$ and given by

$$R_{YX.X} = \frac{R_{YZ} - R_{YX}R_{ZX}}{\{[1-R_{YX}^2][1-R_{ZX}^2]\}^{1/2}} , \tag{1.4}$$

where R_{YZ}, R_{YX} and R_{ZX} are the usual product moment correlation coefficients between Y and Z, Y and X, and Z and X, respectively. That is, if $(X_1,Y_1,Z_1), \ldots, (X_n,Y_n,Z_n)$ denotes a random sample of size n from F, then, for example,

$$R_{YZ} = \frac{\sum_{i=1}^{n} (Y_i-\bar{Y})(Z_i-\bar{Z})}{\{\sum_{i=1}^{n} (Y_i-\bar{Y})^2 \sum_{i=1}^{n} (Z_i-\bar{Z})^2\}^{1/2}} .$$

The intuitive appeal of the statistic $R_{YZ.X}$ arises from the fact that $R_{YZ.X}$ is nothing but the usual product moment correlation coefficient (Pearson's R) calculated from the residuals of the ordinary least squares fit of the linear models given in (1.2). However, a disadvantage of using tests based on $R_{YZ.X}$, which henceforth we shall denote by R_n, is that they all assume that either $E|E'$ or $E'|E$ is normally distributed. These tests may be nonrobust without this assumption, an issue to be investigated in this work.

Another measure for partial correlation, albeit not as popular, is the nonparametric Kendall's partial correlation coefficient given by

$$\tau_{YZ.X} = \frac{\tau_{YZ} - \tau_{YX}\tau_{ZX}}{\{[1-\tau_{YX}^2][1-\tau_{ZX}^2]\}^{1/2}} , \qquad (1.5)$$

where τ_{YZ}, τ_{YX} and τ_{ZX} are the usual Kendall's correlation coefficients (Kendall's tau) calculated on the variables Y and Z, Y and X, and Z and X, respectively. That is, for example,

$$\tau_{YZ} = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{(Y_i - Y_j)(Z_i - Z_j)\} ,$$

where

$$\text{Sgn}\{t\} = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t = 0 \\ -1 & \text{if } t < 0 \end{cases} . \qquad (1.6)$$

Kendall (1962) obtained the statistic $\tau_{YZ.X}$, known also as Kendall's partial tau, as follows. For a fixed ranking of the variable X, he chose two random rankings of the variables Y and Z. For all possible $\binom{n}{2}$ pairs (X_i, Y_i, Z_i) and (X_j, Y_j, Z_j), $i \neq j$, he constructed a 2x2 contingency table in which one category contained the freqencies of agreement (disagreement) of the Y pairs with the X pairs, and the other category contained those of the Z pairs with the X pairs. From this table, Kendall calculated the measure of association commonly known as Kendall's tau-b. Writing the appropriate frequencies in terms of τ_{YZ}, τ_{YX} and τ_{ZX}, he then obtained the partial tau statistic given in (1.5). We have briefly described Kendall's method of obtaining the statistic $\tau_{YZ.X}$ to stress an important fact and that is that $\tau_{YZ.X}$ is <u>not</u> the usual Kendall's tau calculated on the residuals obtained from the linear models (1.2), and that, although $\tau_{YZ.X}$ has the same mathematical structure as $R_{YZ.X}$, it is merely a coincidence.

The lack of popularity of Kendall's partial tau stems from the fact that it has many limitations which are primarily due to its theoretically complex structure. It is not distribution-free, for example, and in fact, it is not even asymptotically distribution-free (its asymptotic variance depends on the underlying distribution of the variable (X,Y,Z)). Magsoodloo (1975) and Magsoodloo and Pallos (1981) have tabulated quantile estimates of a null distribution for $\tau_{YZ.X}$ based on Monte Carlo simulations for a variety of sample sizes. We believe that these quantile estimates are inappropriate for testing conditional independence since they were generated under the hypothesis of "total independence," that is under the assumption that the three variables X, Y and Z are mutually independent. In some preliminary Monte Carlo studies, we used these quantile estimates under the underlying model structure (1.2). As we had expected, the empirical sizes of such tests were highly inflated under the less restrictive hypothesis of conditional independence. For example, for each of 10,000 samples of size n=20 each we have calculated $\tau_{YZ.X}$ for the variables X, Y=X+E and Z=X+E', where the mutually independent standard normal variables X, E and E' were generated by IMSL subroutines. For a nominal α=0.05, each of the 10,000 statistics was compared to the 95th percentile estimates given by Magsoodloo and Pallos (1981). The relative frequency of rejection was found to be 0.138, which indicates that Magsoodloo and Pallos's procedures do not hold their significance levels well under a conditional independence model.

To test the hypothesis of independence of E and E' of (1.2), we propose using Kendall's tau calculated on the residuals. If $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$ and $\hat{\beta}_2$ denote estimates of the regression constants α_1, α_2, β_1

and β_2, respectively, the residuals are

$$U_i = Y_i - \hat{\alpha}_1 - \hat{\beta}_1 X_i$$

and (1.7)

$$V = Z_i - \hat{\alpha}_2 - \hat{\beta}_2 X_i \ , \ i=1,2,\ldots,n \ ,$$

and the test statistic is given by

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{(U_i-U_j)(V_i-V_j)\} \ , \qquad (1.8)$$

with Sgn(t) is as defined in (1.6).

The idea of using Kendall's tau calculated from residuals was considered by Shirahata (1977). In his brief paper, Shirahata tried to show that the difference between a standardized T_n and a standardized S_n converges to zero in probability, where S_n is the usual Kendall's statistic calculated on the variables E and E'. His method of argument is to show via Monte Carlo simulation that, for large n, the correlation between T_n and S_n becomes large while the sample mean of $12(T_n-S_n)^2/\{2n(n-1)(2n+5)\}^{1/2}$ becomes small. From these considerations he concludes that the approximation of T_n to S_n is satisfactory for large n. Randles (1984) also considers applying Kendall's tau to residuals; however, his discussion assumes the X_i's of (1.2) to be known constants rather than random variables as they are considered to be here.

In our study, we compare the performances of tests based on T_n to those based on the Pearson's partial correlation coefficient R_n. The statistic $\tau_{YZ.X}$ will not be included in this study because of the many previously discussed disadvantages associated with it. There are many advantages to using T_n. For example, T_n is asymptotically distribution-free under the model (1.2) and the hypothesis of conditional independence. Further, T_n has many desirable properties regardless of the type of regression parameter estimators used. Also, calculations of asymptotic relative efficiencies (AREs) indicate that, for heavy-tailed distributions and for large n, tests based on T_n have higher relative efficiencies than those based on R_n. These properties will be discussed in detail in chapters 2 and 3. In chapter 2, we discuss the distributional properties of our statistic T_n under the hypothesis of conditional independence, and tabulate the simulated null distribution of T_n when X, E and E' have normal distributions. In chapter 3, we derive an expression for the asymptotic efficiency of T_n relative to R_n [$ARE(T_n,R_n)$], where the class of alternatives of dependence between E and E' is given by the "trivariate reduction" model

$$E = W_1 + \Delta_n W_3$$

and (1.9)

$$E' = W_2 + \Delta_n W_3 ,$$

where W_1, W_2 and W_3 are absolutely continuous and mutually independent random variables and Δ_n is a constant.

In chapter 4, we temporarily turn our attention from the partial correlation problem to a different, yet closely related problem: that of ordinary correlation. Here, the problem of interest is to study the association between the variables Y and Z based on a random sample of pairs $(Y_1,Z_1), \ldots, (Y_n,Z_n)$ from some bivariate continuous distribution F. This problem is commonly known as the test for independence since the available testing procedures based on statistics such as Hoeffding's D, Pearson's R, Spearman's rho and Kendall's tau all test the null hypothesis of independence,

$$H_0: \text{ Y and Z are independent .}$$

Although the hypothesis of independence implies many desirable and convenient theoretical properties, it is our view that, despite its intuitive appeal, such a hypothesis is not broad enough to encompass all situations when no association exists between the variables Y and Z. Suppose, for example, that the pair (Y,Z) has a spherically symmetric distribution with contours of the form given in figure 1.1 (see, for example, Johnson and Ramberg, 1977).

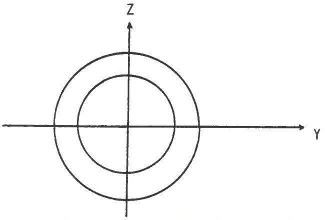

Figure 1.1 Contours of a spherically symmetric distribution

Although Y and Z may be statistically dependent in such cases, they are clearly uncorrelated by all usual definitions of correlation. Moreover, larger values of Y are not associated with larger (or smaller) values of Z, etc. It is situations such as these, when there is no correlation between Y and Z, that we like to include in the null hypothesis. Indeed, some prominent textbooks state their null hypothesis as $\tau=0$, but they calculate the null distribution under independence, not just $\tau=0$, where

$$\tau = P\{(Y_1-Y_2)(Z_1-Z_2)>0\} - P\{(Y_1-Y_2)(Z_1-Z_2)<0\}$$

$$= \text{Probability of concordance} - \text{probability of discordance.} \qquad (1.10)$$

It is our contention that the experimenter often only wishes to detect useful relationships between Y and Z where Y, for example, is useful as a predictor of Z or where, for example, larger Y-values are associated with larger (or smaller) Z-values, etc. Correlation coefficients such as τ attempt to measure these useful relationships.

Of the tests mentioned earlier, Hoeffding's D (see, for example, Hollander and Wolfe, 1973), which is consistent against all types of dependence, is not used as often as Pearson's R or Kendall's tau. This is partly because it is more difficult to compute and interpret, and partly due to its ability to detect all departures from independence, which makes it less powerful at detecting correlated departures. In addition to the fact that the respective consistency classes of the tests based on Pearson's R and Kendall's tau are given by $\rho\neq0$ and $\tau\neq0$, our interest in detecting such alternatives derives from the fact that it is these alternatives that allow us to conclude useful relationships between Y and Z. In view of the above, we would prefer to test the null hypothesis of

$$H_0: \text{ No correlation} \quad \text{versus} \quad H_A: \text{ Correlation ,} \qquad (1.11)$$

viewing this as a test of a non-useful versus a useful relationship between the two variables.

To us, the most natural and intuitive type of correlation is the coefficient τ given in (1.10). The corresponding hypothesis of interest is

$$H_0: \tau = 0 \quad \text{versus} \quad H_A: \tau \neq 0 , \qquad (1.12)$$

or the one-sided alternate hypotheses of positive correlation ($\tau>0$) or negative correlation ($\tau<0$). Note that under H_0, the probability of concordance equals the probability of discordance, so that there is no correlation between Y and Z, in the sense that one variable does not increase or decrease with the other variable. Of course, when Y and Z are independent, $\tau=0$, so that if one rejects the null hypothesis of (1.12), one can safely conclude that the variables Y and Z are indeed dependent, and the dependence is a useful one at least in the sense of predicting direction.

As we mentioned earlier, many authors of statistics textbooks such as Agresti and Agresti (1979) and Ott, Larson and Mendenhall (1983) in testing the hypotheses of (1.12) base their rejection of H_0 on the quantity

$$Z = \frac{\hat{\tau}}{\left[\frac{2(2n+5)}{9n(n-1)} \right]^{1/2}} , \qquad (1.13)$$

where $\hat{\tau}$ is Kendall's estimate of τ given by

$$\hat{\tau} = \frac{1}{\binom{n}{2}} \sum_{i<j} \operatorname{Sgn}\{(Y_i - Y_j)(Z_i - Z_j)\} .$$

We believe that such a test is inappropriate even for large n since the denominator of (1.13) is the standard deviation of $\hat{\tau}$ under the more restrictive hypothesis of independence. Our suspicions of the inappropriateness of such procedures were supported by our Monte Carlo studies where we found that, in some cases when $\tau=0$ but Y and Z are possibly dependent, the empirical α-levels were highly inflated, indicating that this procedure was not maintaining its α-level over the broad class of distributions for which $\tau=0$.

In chapter 4, we review and evaluate the different procedures available for testing (1.12). In particular, we discuss the procedure recommended by Fligner and Rust (1983) and highlight its limitations. Then, we propose a statistic similar to the one given in Fligner and Rust but which has more desirable properties. The performances of all of these procedures are then investigated by a

Monte Carlo study. The results of this study and a summary of conclusions and recommendations are given at the end of chapter 4.

In chapter 5, we return to the partial correlation problem in an effort to investigate the performances of the tests based on the statistics T_n, R_n and some of the statistics studied in chapter 4 but this time applied to the residuals. Through a Monte Carlo study, the empirical powers and sizes of seven different statistics are compared, both under the hypothesis of independence and the hypothesis that $\tau=0$. In each case, the residuals are obtained by two different methods of regression parameter estimation: (i) the ordinary least squares method, and (ii) the method of least absolute regression. The tables of results appear throughout chapter 5 followed by our conclusions and recommendations. A list of related topics for future study appears at the end of chapter 5.

CHAPTER TWO
PROPERTIES OF THE STATISTIC T_n

2.1 Introduction

Let (X_1, Y_1, Z_1), (X_2, Y_2, Z_2), . . . , (X_n, Y_n, Z_n) denote a random sample of observable triples from some absolutely continuous distribution, with distribution function $F(\cdot)$, and let (X, Y, Z) be distributed as (X_1, Y_1, Z_1). To test the conditional independence of Y and Z, holding X constant, we shall assume that each of Y and Z is linearly related to X as follows,

$$Y_i = \alpha_1 + \beta_1 X_i + E_i$$

and

$$Z_i = \alpha_2 + \beta_2 X_i + E_i' \quad , \; i = 1, 2, \ldots, n \; , \tag{2.1.1}$$

where α_1, β_1, α_2 and β_2 are unknown parameters which need to be estimated. Here, $X_1, X_2, \ldots, X_n$, which will be referred to as the "covariate terms," are independent identically distributed (i.i.d.) random variables with an absolutely continuous distribution function $F_X(\cdot)$, mean μ_X and variance σ_X^2. The "error terms" (E_i, E_i'), $i = 1, 2, \ldots, n$, are i.i.d. absolutely continuous bivariate random variables. The respective marginal distribution of E_i (E_i') is assumed to have mean zero, distribution function $H_1(\cdot)$ $(H_2(\cdot))$ and variance $\sigma_E^2(\sigma_{E'}^2)$. Further, it will be assumed that X_i is independent of (E_i, E_i'), $i = 1, 2, \ldots, n$.

The hypothesis of interest is

$$H_0: \; E_i \text{ and } E_i' \text{ are independent, } i = 1, 2, \ldots, n.$$

Our proposed test statistic, T_n, is the Kendall's tau statistic applied to the residuals (i.e., to the estimates of the unobservable error terms, $E_1, E_2, \ldots, E_n$ and $E_1', E_2', \ldots, E_n'$). If $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$, and $\hat{\beta}_2$ denote the estimates of α_1, α_2, β_1, and β_2, respectively, the residuals are given by

$$U_i = Y_i - \hat{\alpha}_1 - \hat{\beta}_1 X_i$$

and $$V_i = Z_i - \hat{\alpha}_2 - \hat{\beta}_2 X_i \quad , \; i = 1, 2, \ldots, n, \qquad (2.1.2)$$

and the proposed test statistic is

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}[(U_i - U_j)(V_i - V_j)], \qquad (2.1.3)$$

where

$$\mathrm{Sgn}(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t = 0 \\ -1 & \text{if } t < 0 \end{cases} .$$

In the sections to follow, we shall discuss the properties of this statistic. In section 2.2 it will be shown that the distribution of T_n is free of the regression constants α_1, α_2, β_1, and β_2, the location parameter μ_X, and the scale parameters σ_X^2, σ_E^2 and $\sigma_{E'}^2$, provided that the estimates of the regression constants $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$, and $\hat{\beta}_2$ satisfy certain general properties. In section 2.3 the small sample moments and the symmetric distribution of T_n under the null hypothesis of independence will be discussed. In section 2.4 the asymptotic distribution of T_n under H_0 will be developed. Section 2.5 will contain the tables of the small sample null distribution of the T_n statistic as generated by a Monte Carlo simulation study when the X_i's, E_i's, and E_i''s are normally distributed.

2.2 The Effects of Parameters on the Distribution of T_n

Unlike the usual Kendall's tau statistic, T_n is not a distribution-free statistic even under the hypothesis of the independence of the "error terms." Its distribution depends on the distribution of the X_i's, E_i's and E_i''s, i = 1, 2, . . . , n. To see this, we write the residuals given in (2.1.2) in terms of the error terms to obtain

$$U_i = Y_i - \hat{\alpha}_1 - \hat{\beta}_1 X_i$$

$$= (\alpha_1 - \hat{\alpha}_1) - (\hat{\beta}_1 - \beta_1) X_i + E_i ,$$

and similarly, (2.2.1)

$$V_i = (\alpha_2 - \hat{\alpha}_2) - (\hat{\beta}_2 - \beta_2) X_i + E_i' , \quad i = 1, 2, \ldots, n.$$

The statistic T_n is the Kendall's correlation coefficient (Kendall's tau) calculated on the pairs

$$(U_i, V_i) = [(\alpha_1 - \hat{\alpha}_1) - (\hat{\beta}_1 - \beta_1) X_i + E_i \ , \ (\alpha_2 - \hat{\alpha}_2) - (\hat{\beta}_2 - \beta_2) X_i + E_i'], \qquad (2.2.2)$$

i = 1, 2,..., n.

The distribution-free property of the usual Kendall's statistic under H_0 results from the fact that under the hypothesis of independence the two elements of the pair are exchangeable, and there is independence between pairs. However, in the set-up considered here, the two elements of the pair are not exchangeable due to the presence of the X_i's in both elements. (Note that in (2.2.1) the X_i's appear both explicitly and implicitly through the estimators $\hat{\alpha}_1$, $\hat{\alpha}_2$, $\hat{\beta}_1$, and $\hat{\beta}_2$).

Although the statistic T_n is not distribution-free, its distribution does not depend on the parameters α_1, α_2, β_1, β_2, μ_X, σ_X^2, σ_E^2 and $\sigma_{E'}^2$ under "translation" and "scale" properties to be discussed later.

The statistic T_n is free of the terms α_1, α_2, $\hat{\alpha}_1$ and $\hat{\alpha}_2$, since these quantities are cancelled out by taking the differences of the residuals. Writing

$$\text{Sgn}\ \{(U_i-U_j)(V_i-V_j)\}$$

$$= \text{Sgn}\ \{[(E_i-E_j)-(\hat{\beta}_1-\beta_1)(X_i-X_j)][(E_i'-E_j')-(\hat{\beta}_2-\beta_2)(X_i-X_j)]\},$$

we see that

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{[(E_i-E_j)-(\hat{\beta}_1-\beta_1)(X_i-X_j)][(E_i'-E_j')-(\hat{\beta}_2-\beta_2)(X_i-X_j)]\}. \tag{2.2.3}$$

Thus, without loss of generality, the intercept terms α_1 and α_2 may be taken to be zero. Furthermore, the distribution of T_n is free of the location parameter μ_X. For, if $\mu_X \neq 0$, consider the transformed zero-mean random variables $X_i^* = X_i - \mu_X$, $i = 1, 2, \ldots, n$. The underlying model may now be written as

$$Y_i = \alpha_1 + \beta_1(X_i^* + \mu_X) + E_i$$

$$= \alpha_1' + \beta_1 X_i^* + E_i \ ,$$

and

$$Z_i = \alpha_2' + \beta_2 X_i^* + E_i' \ ,$$

$i = 1, 2, \ldots, n$,

where $\qquad \alpha_1' = \alpha_1 + \beta_1\mu_X$ and $\alpha_2' = \alpha_2 + \beta_2\mu_X$.

By the preceding argument, T_n is free of α_1' and α_2', and is therefore free of the location parameter μ_X.

To ensure that the distribution of T_n is free of the remaining parameters β_1, β_2, σ_X^2, σ_E^2 and $\sigma_{E'}^2$, it is sufficient that the slope estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ satisfy the following properties.

<u>"Translation" property 2.2.4</u>

Assume each $\hat{\beta}_i$, $i = 1, 2$, satisfies

$$\hat{\beta}_i(x_1,\ldots,x_n;\ y_1+cx_1,\ldots,\ y_n+cx_n) = \hat{\beta}_i(x_1,\ldots,x_n;\ y_1,\ldots,y_n) + c$$

for every $x_1, \ldots, x_n, y_1, \ldots, y_n$ and c.

<u>"Scale" property 2.2.5</u>

Assume each $\hat{\beta}_i$, $i = 1, 2$, satisfies

$$\hat{\beta}_i(ax_1,\ldots,ax_n;\ by_1,\ldots,\ by_n) = \frac{b}{a}\,\hat{\beta}_i(x_1,\ldots,x_n;\ y_1,\ldots,\ y_n)$$

for every $x_1, \ldots, x_n, y_1, \ldots, y_n$, b and $a \neq 0$.

From expression (2.2.3), we can see that the statistic T_n involves quantities of the form

$$E_i - (\hat{\beta}_1 - \beta_1)\, X_i$$

and (2.2.6)

$$E_i' - (\hat{\beta}_2 - \beta_2)\, X_i \quad .$$

Applying property 2.2.4 with $c = -\beta_1$ and $c = -\beta_2$, respectively, we obtain

$$\hat{\beta}_1(x_1,\ldots,x_n;\, y_1,\ldots,y_n)-\beta_1 = \hat{\beta}_1(x_1,\ldots,x_n;\, y_1-\beta_1x_1,\ldots,y_n-\beta_1x_n)$$

and

$$\hat{\beta}_2(x_1,\ldots,x_n;\, z_1,\ldots,z_n)-\beta_2 = \hat{\beta}_2(x_1,\ldots,x_n;\, z_1-\beta_2x_1,\ldots,z_n-\beta_1x_n)$$

so that the quantities $(\hat{\beta}_1 - \beta_1)$ and $(\hat{\beta}_2 - \beta_2)$ may be replaced by

$$\hat{\beta}_1^* = \hat{\beta}_1\ (X_1,\ldots,X_n;\, Y_1-\beta_1X_1,\ldots,Y_n-\beta_1X_n)$$

and

$$\hat{\beta}_2^* = \hat{\beta}_2\ (X_1,\ldots,X_n;\, Z_1-\beta_2X_1,\ldots,Z_n-\beta_2X_n)\ ,$$

without changing the value of T_n. These new estimators $\hat{\beta}_1^*$ and $\hat{\beta}_2^*$ are the slope estimators obtained by replacing Y_i by $Y_i-\beta_iX_i$ and Z_i by $Z_i-\beta_2X_i$, respectively, in the model structure

$$Y_i = \alpha_1 + \beta_1X_i + E_i$$

and

$$Z_i = \alpha_2 + \beta_2X_i + E_i'\ ,\ i = 1, 2,\ldots, n\ .$$

This is equivalent to using the slope estimators obtained from the models

$$Y_i = \alpha_1 + E_i \text{ and } Z_i = \alpha_2 + E_i' \text{ , } i = 1, 2, \ldots, n \text{ ,}$$

which is the usual model with $\beta_1 = \beta_2 = 0$. Consequently, the statistic T_n does not depend on the values of the slope parameters β_1 and β_2.

Next, "scale" property 2.2.5 is used to show the distribution of T_n is free of the scale parameters σ_X^2, σ_E^2 and $\sigma_{E'}^2$. The statistic T_n involves residuals of the form

$$U_i = Y_i - \hat{\beta}_1 X_i$$

and

$$V_i = Z_i - \hat{\beta}_2 X_i \text{ , } \quad i = 1, 2, \ldots, n \text{ .}$$

From property 2.2.5 with $a = 1/\sigma_X$ and $b = 1$,

$$\sigma_X \hat{\beta}_i(x_1,\ldots,x_n;\ y_1,\ldots,y_n) = \hat{\beta}_i(x_1/\sigma_X,\ldots,x_n/\sigma_X;\ y_1,\ldots,y_n) \text{ ,}$$

$i = 1, 2$, so that the residual estimates above may be written as

$$U_i = Y_i - \sigma_X \hat{\beta}_1 \ (X_1,\ldots,X_n;\ Y_1,\ldots,Y_n) \ (\frac{X_i}{\sigma_X})$$

$$= Y_i - \hat{\beta}_1 \ (X_1/\sigma_X,\ldots,X_n/\sigma_X;\ Y_1,\ldots,Y_n) \ (\frac{X_i}{\sigma_X}) \text{ ,}$$

and

$$V_i = Y_i - \hat{\beta}_2 \ (X_1/\sigma_X,\ldots,X_n/\sigma_X;\ Z_1,\ldots,Z_n) \ (\frac{X_i}{\sigma_X}) \text{ ,}$$

which indicates that the X_i's may be replaced by their standardized forms, X_i/σ_X, without changing the values of the residuals. Thus, T_n is free of the scale parameter of the X's.

From (2.2.6) and the discussion immediately following, we can see that T_n may be written in terms of residual estimates of the form

$$R_i = E_i - \hat{\beta}_1\,(X_1,\ldots,X_n;\ E_1,\ldots,E_n)\ X_i\ ,$$

and

$$R_i' = E_i' - \hat{\beta}_2\,(X_1,\ldots,X_n;\ E_1',\ldots,E_n')\ X_i\ . \qquad (2.2.7)$$

Applying "scale" property 2.2.5 with a=1 and $b=\sigma_1$, we have

$$\hat{\beta}_1(X_1,\ldots,X_n;\ E_1/\sigma_E,\ldots,E_n/\sigma_E) = \frac{1}{\sigma_E}\,\hat{\beta}_1(X_1,\ldots,X_n;\ E_1,\ldots,E_n).$$

Similarly,

$$\hat{\beta}_2(X_1,\ldots,X_n;\ E_1'/\sigma_{E'},\ldots,E_n'/\sigma_{E'}) = \frac{1}{\sigma_{E'}}\,\hat{\beta}_2(X_1,\ldots,X_n;\ E_1',\ldots,E_n')\ ,$$

so that replacing the error terms E_i (E_i') in (2.2.7) by their standardized forms E_i/σ_E ($E_i'/\sigma_{E'}$), i = 1, 2, . . . , n, will result in transforming the residual estimates to R_i/σ_E ($R_i'/\sigma_{E'}$). However, the T_n statistic which is based on the sign of the product of the residual differences is not affected by such scaling, since both σ_E and $\sigma_{E'}$ are positive constants, and hence the distribution of T_n is free of these scale parameters.

Thus far in this section we have shown that if $\hat{\beta}_1$ and $\hat{\beta}_2$ satisfy properties 2.2.4 and 2.2.5, the distribution of the statistic T_n is independent of the regression constants used in the linear models, the location and scale of the "covariate term" X, and of the scale parameters of the error terms. In the remainder of this section, we shall demonstrate that properties 2.2.4 and 2.2.5 are very natural, and that the three types of slope estimators we have used in this study, namely the least squares estimator (OLS), the least absolute value estimator (LAV), and Theil's slope estimator, all satisfy these properties under the linear model

$$Y_i = \alpha + \beta X_i + E_i \ , \ i = 1, 2, \ldots, n \ .$$

The least square (OLS) estimator of the slope is given by

$$\hat{\beta}(x_1,\ldots,x_n;\ y_1,\ldots,y_n) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} \ ,$$

so that

$$\text{(i)}\quad \hat{\beta}(x_1,\ldots,x_n;\ y_1+cx_1,\ldots,y_n+cx_n) = \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i+cx_i-\bar{y}-c\bar{x})}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$= \frac{\sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{\sum_{i=1}^{n}(x_i-\bar{x})^2} + \frac{c\sum_{i=1}^{n}(x_i-\bar{x})^2}{\sum_{i=1}^{n}(x_i-\bar{x})^2}$$

$$= \hat{\beta}(x_1,\ldots,x_n;\ y_1,\ldots,y_n) + c \ ,$$

$$\text{(ii)}\quad \hat{\beta}(ax_1,\ldots,ax_n;\ by_1,\ldots,by_n) = \frac{\sum_{i=1}^{n}(ax_i-a\bar{x})(by_i-b\bar{y})}{\sum_{i=1}^{n}(ax_i-a\bar{x})^2}$$

$$= \frac{a\ b\ \sum_{i=1}^{n}(x_i-\bar{x})(y_i-\bar{y})}{a^2\ \sum_{i=1}^{n}(x_1-\bar{x})^2}$$

$$= \frac{b}{a}\,\hat{\beta}(x_1,\ldots,x_n;\ y_1,\ldots,y_n)\ ,$$

provide $a \neq 0$, and 2.2.4 and 2.2.5 are satisfied.

The least absolute value (LAV) estimators of α and β, denoted by $\hat{\alpha}$ and $\hat{\beta}$ respectively, are the values of α and β which satisfy

$$\min_{\alpha,\beta} \sum_{i=1}^{n} |\ y_i-\alpha-\beta x_i\ | = \sum_{i=1}^{n} |\ y_i-\hat{\alpha}-\hat{\beta}x_i\ | \ . \tag{2.2.8}$$

To see that the LAV slope estimator, $\hat{\beta}$, satisfies properties 2.2.4 and 2.2.5, note that

$$\begin{aligned}\min_{\alpha,\beta} \sum_{i=1}^{n} |\ y_i-\alpha-\beta x_i\ | &= \sum_{i=1}^{n} |\ y_i-\hat{\alpha}-\hat{\beta}x_i\ | \\ &= \sum_{i=1}^{n} |\ y_i-\hat{\alpha}-(\hat{\beta}+c-c)x_i\ | \\ &= \sum_{i=1}^{n} |\ y_i+cx_i-\hat{\alpha}-(\hat{\beta}+c)x_i\ |\end{aligned}$$

so that if y_i is replaced by y_i+cx_i, $i = 1, 2, \ldots, n$, the new slope estimate is given by $(\hat{\beta}+c)$, which proves the "translation" property 2.2.4. Also, for $a \neq 0$ and $b \neq 0$,

$$\min_{\alpha,\beta} \sum_{i=1}^{n} | y_i-\alpha-\beta x_i | = \sum_{i=1}^{n} | y_i-\hat{\alpha}-\hat{\beta}x_i |$$

$$= \sum_{i=1}^{n} | \frac{b}{b} y_i-\hat{\alpha}-\hat{\beta}(\frac{ax_i}{a}) |$$

$$= \frac{1}{|b|} \sum_{i=1}^{n} | by_i-b\hat{\alpha}- \frac{b}{a} \hat{\beta}(ax_i) |$$

This last expression indicates that when x_i is replaced by ax_i, and y_i is replaced by by_i, $i = 1,2, \ldots , n$, the new slope estimate is given by $\frac{b}{a} \hat{\beta}$. Note also that b=0 is equivalent to all the y_i's equal to zero, in which case the LAV estimates are $\hat{\alpha}=0$ and $\hat{\beta}=0$. This proves the "scale" property 2.2.5.

Theil's estimate of the slope (see Sen, 1968) is the median of the $\binom{n}{2}$ slopes obtained from the (X_i, Y_i) pairs, i.e.,

$$\hat{\beta}(x_1,\ldots,x_n; y_1,\ldots,y_n) = \underset{i<j}{\text{median}} \left\{ \frac{y_j-y_i}{x_j-x_i} \right\} .$$

We assume that all the x_i's are distinct because they have a continuous distribution. It follows that

(i) $$\hat{\beta}(x_1,\ldots,x_n;\ y_1+cx_1,\ldots,y_n+cx_n) = \underset{i<j}{\text{median}} \left\{ \frac{y_j-cx_j-y_i-cx_i}{x_j-x_i} \right\}$$

$$= \underset{i<j}{\text{median}} \left\{ \frac{y_j-y_i}{x_j-x_i} + c \right\}$$

$$= \hat{\beta}(x_1,\ldots,x_n;\ y_1,\ldots,y_n) + c ,$$

and

(ii) $$\hat{\beta}(ax_1,\ldots,ax_n;\ bx_1,\ldots,bx_n) = \underset{i<j}{\text{median}} \left\{ \frac{by_j-by_i}{ax_j-ax_i} \right\}$$

$$= \frac{b}{a}\,\hat{\beta}(x_1,\ldots,x_n;\ y_1,\ldots,y_n) ,$$

and therefore this estimator also satisfies the properties 2.2.4 and 2.2.5.

2.3 The Null Hypothesis Distribution of T_n

The test statistic T_n given by

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{[(Y_i-Y_j)-\hat{\beta}_1(X_i-X_j)][(Z_i-Z_j)-\hat{\beta}_2(X_i-X_j)]\} \qquad (2.3.1)$$

$$= \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{[(E_i-E_j)-(\hat{\beta}_1-\beta_1)(X_i-X_j)][(E_i'-E_j')-(\hat{\beta}_2-\beta_2)(X_i-X_j)]\} \qquad (2.3.2)$$

would be a U-statistic of degree 2, except for the presence of the terms $\hat{\beta}_1$ and $\hat{\beta}_2$. The symmetric kernel of this U-statistic with these two auxiliary estimators in the kernel is

$$h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \hat{\beta}_1,\hat{\beta}_2) =$$

$$\text{Sgn}\{[(E_1-E_2)-(\hat{\beta}_1-\beta_1)(X_1-X_2)][(E_1'-E_2')-(\hat{\beta}_2-\beta_2)(X_1-X_2)]\}\ ,$$

where $\underset{\sim}{S}_i = (X_i, E_i, E_i')$, $i = 1, 2$. Because this U-statistic involves the estimated parameters $\hat{\beta}_1$ and $\hat{\beta}_2$, ordinary U-statistic theorems (see, for example, Randles and Wolfe, 1979) cannot be used to develop its large sample distributional properties. In what follows, we will use equal-in-distribution arguments to show that under H_0 when the distribution of at least one of the error terms, say E, is symmetric about zero, T_n is symmetric about its mean zero. In addition, we shall derive an expression for $\text{Var}[T_n]$, and discuss the null asymptotic distribution of T_n using a theorem by Randles (1982).

Since the distribution of T_n is free of the parameters α_1, α_2, β_1 and β_2, with no loss of generality we assume each of them to be zero in the following discussion. The statistic may be written as

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\ \{[(E_i-E_j)-\hat{\beta}_1(X_i-X_j)][(E_i'-E_j')-\hat{\beta}_2(X_i-X_j)]\}\ , \qquad (2.3.3)$$

where $\hat{\beta}_1$ ($\hat{\beta}_2$) is a function of the X_i's and the E_i's (E_i''s). Let

$$Q_{ij} = [(E_i-E_j) - \hat{\beta}_1(X_i-X_j)]$$

and

$$Q_{ij}' = [(E_i'-E_j') - \hat{\beta}_2(X_i-X_j)]$$

and write

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{\, Q_{ij}Q'_{ij} \,\} \;.$$

Now suppose the distribution of the E_i's is symmetric about zero, that is $E_i \overset{d}{=} -E_i$, $i = 1, 2, \ldots, n$. From the "scale" property 2.2.5, we note that

$$\hat{\beta}_1(X_1,\ldots,X_n;\, -E_1,\ldots,-E_n) = -\hat{\beta}_1(X_1,\ldots,X_n;\, E_1,\ldots,\, E_n) \;.$$

Using the independence of the E_i's and their independence from the E_i''s and X_i's, we have

$$(X_1,E_1,E_1',\ldots,X_n,E_n,E_n') \overset{d}{=} (X_1,-E_1,E_1',\ldots,X_n,-E_n,E_n') \;.$$

Computing T_n on both sides of the equal in distribution sign yields

$$\begin{aligned} T_n &= \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{\, Q_{ij}Q'_{ij} \,\} \\ &\overset{d}{=} \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{\, -Q_{ij}Q'_{ij} \,\} \\ &= -\,T_n \;, \end{aligned}$$

and, therefore, under H_0 and the assumption of the symmetry of one set of error terms, T_n is symmetric about its mean of zero. Note that when the assumption of symmetry is dropped $E[T_n]$ is, in general, different from zero even under H_0, and will be given by

$$E[T_n] = P\{Q_{12}Q'_{12}>0\} - P\{Q_{12}Q'_{12}<0\}$$

$$= 2P\{Q_{12}>0,\ Q'_{12}>0\} + 2P\{Q_{12}<0,\ Q'_{12}<0\} - 1\ . \qquad (2.3.4)$$

The expression for the null variance of T_n is rather complex since the distribution of T_n depends on the underlying distributions of the X_i's, E_i's and E'_i's, and the type of slope estimators $\hat{\beta}_1$ and $\hat{\beta}_2$ used to generate the residuals. This, however, causes no limitation to the applicability of our results for large sample sizes as we shall demonstrate later, since the limiting null variance is free of the underlying distributions and the kind of slope estimators used. For the sake of completeness, however, we will include the general form for the null variance of T_n. We write

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} h(\underset{\sim}{S}_i, \underset{\sim}{S}_j;\ \underset{\sim}{\hat{\beta}})\ ,$$

where

$$h(\underset{\sim}{S}_i, \underset{\sim}{S}_j;\ \underset{\sim}{\hat{\beta}}) = \mathrm{Sgn}\ \{[(E_i-E_j)-\hat{\beta}_1(X_i-X_j)][(E'_i-E'_j)-\hat{\beta}_2(X_i-X_j)]\}$$

$$= \mathrm{Sgn}\ \{Q_{ij}Q'_{ij}\}$$

and $\underset{\sim}{\hat{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$. Let θ denote the mean of T_n given in (2.3.4). That is,

$$\theta = 2P\{Q_{12}>0,\ Q'_{12}>0\} + 2P\{Q_{12}<0,\ Q'_{12}<0\} - 1\ .$$

Then,

$$\mathrm{Var}[T_n] = E[\{\frac{1}{\binom{n}{2}} \sum_{i<j} [h(\underset{\sim}{S}_i,\underset{\sim}{S}_j;\ \hat{\underset{\sim}{\beta}})-\theta]\}^2]$$

$$= \frac{1}{\binom{n}{2}^2} \sum_{i<j} \sum_{i'<j'} E[\{h(\underset{\sim}{S}_i,\underset{\sim}{S}_j;\ \hat{\underset{\sim}{\beta}})-\theta\}\{h(\underset{\sim}{S}_{i'},\underset{\sim}{S}_{j'};\ \hat{\underset{\sim}{\beta}})-\theta\}]$$

$$= \frac{1}{\binom{n}{2}^2} \sum_{i<j} \sum_{i'<j'} \mathrm{Cov}[h(\underset{\sim}{S}_i,\underset{\sim}{S}_j;\ \hat{\underset{\sim}{\beta}}),\ h(\underset{\sim}{S}_{i'},\underset{\sim}{S}_{j'};\ \hat{\underset{\sim}{\beta}})]\ .$$

There are three types of terms in the above expression:

Type 0, where the two kernels involve no subscripts in common. There are

$$\binom{n}{2}\binom{2}{0}\binom{n-2}{2} = \binom{n}{2}\frac{(n-2)(n-3)}{2}$$

such terms.

Type 1, are terms with one subscript in common. There are

$$\binom{n}{2}\binom{2}{1}\binom{n-2}{1} = 2\binom{n}{2}(n-2)$$

such terms.

Type 2 terms have two subscripts in common. There are $\binom{n}{2}$ of them.

Denoting the expectations of such terms by ζ_0, ζ_1 and ζ_2, respectively, we have

$$\mathrm{Var}[T_n] = \frac{1}{\binom{n}{2}} [\frac{(n-2)(n-3)}{2} \zeta_0 + 2(n-2)\ \zeta_1 + \zeta_2]\ , \qquad (2.3.5)$$

where

$$\zeta_0 = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \hat{\underset{\sim}{\beta}}),\ h(\underset{\sim}{S}_3,\underset{\sim}{S}_4;\ \hat{\underset{\sim}{\beta}})] - \theta^2$$

$$= E[\text{Sgn}\ \{Q_{12}Q'_{12}\}\ \text{Sgn}\ \{Q_{34}Q'_{34}\}] - \theta^2$$

$$= P\{Q_{12}Q'_{12}>0, Q_{34}Q'_{34}>0\} + P\{Q_{12}Q'_{12}<0, Q_{34}Q'_{34}<0\}$$

$$-P\{Q_{12}Q'_{12}>0, Q_{34}Q'_{34}<0\} - P\{Q_{12}Q'_{12}<0, Q_{34}Q'_{34}>0\} - \theta^2$$

$$= 2P\{Q_{12}Q'_{12}>0, Q_{34}Q'_{34}>0\} + 2P\{Q_{12}Q'_{12}<0, Q_{34}Q'_{34}<0\} - 1 - \theta^2\ , \qquad (2.3.6)$$

$$\zeta_1 = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \hat{\underset{\sim}{\beta}})\ h(\underset{\sim}{S}_1,\underset{\sim}{S}_3;\ \hat{\underset{\sim}{\beta}})] - \theta^2$$

$$= 2P\{Q_{12}Q'_{12}>0, Q_{13}Q'_{13}>0\} + 2P\{Q_{12}Q'_{12}<0, Q_{13}Q'_{13}<0\} - 1 - \theta^2\ , \qquad (2.3.7)$$

and

$$\zeta_2 = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \hat{\underset{\sim}{\beta}})\ h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \hat{\underset{\sim}{\beta}})] - \theta^2$$

$$= 1 - \theta^2\ , \text{ if } n > 2. \qquad (2.3.8)$$

We have shown earlier that when the error terms of one of the underlying linear models, the E's say, are symmetrically distributed about zero, then $\theta = 0$ under H_0. However, this does not significantly simplify the expressions for ζ_0 and ζ_1, since to evaluate these expressions one needs to know the distribution of the covariate term, X, and the joint distributions of variables of the form $\{Q_{ij}, Q_{kl}\}$ and $\{Q'_{ij}, Q'_{kl}\}$. We shall demonstrate this by calculating the null hypothesis value of $\text{Var}[T_n]$ in the special case when E_i (E'_i), $i = 1, 2, \ldots, n$, have the standard normal distribution, and when $\hat{\beta}_1$ and $\hat{\beta}_2$ are the ordinary least squares estimators of β_1 and β_2, respectively. Under the symmetry of the error terms, $\theta = 0$ and conditional on $\underset{\sim}{X} = \underset{\sim}{x}$

$$[(E_1-E_2)-\hat{\beta}_1(x_1-x_2),\ (E_3-E_4)-\hat{\beta}_1(x_3-x_4)]$$

$$\overset{d}{=} [-(E_1-E_2)+\hat{\beta}_1(x_1-x_2),\ -(E_3-E_4)+\hat{\beta}_1(x_3-x_4)] \ ,$$

since by property 2.2.5 of $\hat{\beta}_1$

$$\hat{\beta}_1(x_1,\ldots,x_n;\ -E_1,\ldots,-E_n) = -\hat{\beta}_1(x_1,\ldots,x_n;\ E_1,\ldots,E_n).$$

A similar statement can be made for the terms involving E'. Taking expectations with respect to $\underset{\sim}{X}$, and using the above arguments and the null hypothesis of the independence of E_i and E_i', the expression for ζ_0 given in (2.3.6) simplifies to

$$\zeta_0 = E_{\underset{\sim}{X}}\{8P[Q_{12}>0,Q_{34}>0]P[Q_{12}'>0,Q_{34}'>0]$$

$$+8P[Q_{12}>0,Q_{34}<0]P[Q_{12}'>0,Q_{34}'<0] - 1 \mid \underset{\sim}{X} = \underset{\sim}{x}\} \ . \qquad (2.3.9)$$

Similarly,

$$\zeta_1 = E_{\underset{\sim}{X}}\{8P[Q_{12}>0,Q_{13}>0]P[Q_{12}'>0,Q_{13}'>0]$$

$$+8P[Q_{12}>0,Q_{13}<0]P[Q_{12}'>0,Q_{13}'<0] - 1 \mid \underset{\sim}{X} = \underset{\sim}{x}\} \ , \qquad (2.3.10)$$

where $E_{\underset{\sim}{X}}$ denotes expectation with respect to the vector of covariates, $\underset{\sim}{X}$. To obtain an expression for $\text{Var}[T_n]$, we need the following lemma which we will state without proof:

Lemma 2.3.11

Suppose $\binom{X}{Y} \sim \text{BVN}\left[\binom{0}{0}, \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}\right]$,

then, $P\,[X \geq 0 \,,\, Y \geq 0] = \frac{1}{4} + \frac{1}{2\pi}\,\text{Sin}^{-1}(\rho)$,

and consequently, $P\,[X \geq 0,\, Y < 0] = \frac{1}{4} - \frac{1}{2\pi}\,\text{Sin}^{-1}(\rho)$

(see for example, Cramér, 1966, p. 290).

The least squares estimator $\hat{\beta}_1$ is given by

$$\hat{\beta}_1 = \frac{\sum_{i=1}^{n} (X_i-\bar{X})E_i}{\sum_{i=1}^{n} (X_i-\bar{X})^2} = \frac{\sum_{i=1}^{n} (X_i-\bar{X})E_i}{S_{XX}} ,$$

where

$$S_{XX} = \sum_{i=1}^{n} (X_i-\bar{X})^2 ,$$

so that, for example,

$$P\{Q_{12}>0, Q_{13}>0 \mid \underset{\sim}{X} = \underset{\sim}{x}\}$$

$$= P\{[(E_1-E_2) - \frac{(x_1-x_2)}{S_{xx}} \sum_{i=1}^{n} (x_i-\bar{x})E_i]>0,$$

$$[(E_1-E_3) - \frac{(x_1-x_3)}{S_{xx}} \sum_{i=1}^{n} (x_i-\bar{x})E_i]>0\}$$

$$\equiv P\{q_{12}>0, q_{13}>0\} ,$$

where

$$q_{ij} = (E_i-E_j) - \frac{(x_i-x_j)}{S_{xx}} \sum_{k=1}^{n} (x_k-\bar{x})E_k .$$

This probability statement involves linear combinations of i.i.d. standard normal random variables. The combinations are also zero mean normal variables, so that by lemma 2.3.11

$$P\{q_{12}>0, q_{13}>0\} = \frac{1}{4} + \frac{1}{2\pi} \mathrm{Sin}^{-1}[\rho(\underset{\sim}{x},12,13)] ,$$

where

$$\rho(\underset{\sim}{x},12,13) = \frac{\mathrm{Cov}(q_{12},q_{13})}{\{\mathrm{Var}(q_{12})\ \mathrm{Var}(q_{13})\}^{1/2}} .$$

But,

$$\mathrm{Cov}(q_{12},q_{13}) = 1 - \frac{(x_1-x_3)(x_1-\bar{x})}{S_{xx}} + \frac{(x_1-x_3)(x_2-\bar{x})}{S_{xx}}$$

$$- \frac{(x_1-x_2)(x_1-\bar{x})}{S_{xx}} + \frac{(x_1-x_2)(x_3-\bar{x})}{S_{xx}} + \frac{(x_1-x_2)(x_1-x_3)}{S_{xx}}$$

$$= 1 - \frac{(x_1-x_2)(x_1-x_3)}{S_{xx}} ,$$

$$\mathrm{Var}(q_{12}) = 1 + 1 + \frac{(x_1-x_2)^2}{S_{xx}} - \frac{2(x_1-x_2)(x_1-\bar{x})}{S_{xx}} + \frac{2(x_1-x_2)(x_2-\bar{x})}{S_{xx}}$$

$$= 2 - (x_1-x_2)^2/S_{xx} ,$$

and

$$\mathrm{Var}(q_{13}) = 2 - (x_1-x_3)^2/S_{xx} .$$

These yield

$$\rho(\underset{\sim}{x},12,13) = \frac{S_{xx}-(x_1-x_2)(x_1-x_3)}{\{[2S_{xx}-(x_1-x_2)^2][2S_{xx}-(x_1-x_3)^2]\}^{1/2}} . \tag{2.3.12}$$

To evaluate $P\{q_{12}>0,q_{34}>0\}$, we need the quantity

$$\rho(\underset{\sim}{x},12,34) = \frac{\mathrm{Cov}(q_{12},q_{34})}{\{\mathrm{Var}(q_{12})\ \mathrm{Var}(q_{34})\}^{1/2}} ,$$

which may similarly be calculated to be

$$\rho(\underset{\sim}{x},12,34) = \frac{-(x_1-x_2)(x_3-x_4)}{\{[2S_{xx}-(x_1-x_2)^2][2S_{xx}-(x_3-x_4)^2\}^{1/2}} . \tag{2.3.13}$$

The probabilities associated with q'_{ij} involve the same quantities given above, since $E_i \overset{d}{=} E'_i$, i = 1, 2, . . . , n. From the expressions for ζ_0 and ζ_1 given in (2.3.9) and (2.3.10) we obtain,

$$\zeta_0 = E_{\underset{\sim}{X}}[8\{\frac{1}{4} + \frac{1}{2\pi}\operatorname{Sin}^{-1}[\rho(\underset{\sim}{x},12,34)]\}^2$$

$$+8\{\frac{1}{4} - \frac{1}{2\pi}\operatorname{Sin}^{-1}[\rho(\underset{\sim}{x},12,34)]\}^2 - 1 \mid \underset{\sim}{X} = \underset{\sim}{x}]$$

$$= 4E_{\underset{\sim}{X}}[\{\operatorname{Sin}^{-1}[\rho(\underset{\sim}{x},12,34)]\}^2 \mid \underset{\sim}{X} = \underset{\sim}{x}]/\pi^2 ,$$

and

$$\zeta_1 = 4E_{\underset{\sim}{X}}[\{\operatorname{Sin}^{-1}[\rho(\underset{\sim}{x},12,13)]\}^2 \mid \underset{\sim}{X} = \underset{\sim}{x}]/\pi^2 .$$

From (2.3.5), we get

$$\operatorname{Var}[T_n] = \frac{1}{\binom{n}{2}}[\frac{(n-2)(n-3)}{2}\zeta_0 + 2(n-2)\zeta_1 + 1]$$

where

$$\zeta_0 = \frac{4}{\pi^2}E_{\underset{\sim}{X}}\left[\left\{\operatorname{Sin}^{-1}\left[\frac{-(X_1-X_2)(X_3-X_4)}{\{[2S_{XX}-(X_1-X_2)^2][2S_{XX}-(X_3-X_4)^2]\}^{1/2}}\right]\right\}^2\right]$$

and

$$\zeta_1 = \frac{4}{\pi^2}E_{\underset{\sim}{X}}\left[\left\{\operatorname{Sin}^{-1}\left[\frac{S_{XX}-(X_1-X_2)(X_3-X_4)}{\{[2S_{XX}-(X_1-X_2)^2][2S_{XX}-(X_1-X_3)^2]\}^{1/2}}\right]\right\}^2\right]$$

with $E_{\underset{\sim}{X}}$ indicating expectation with respect to the random vector $\underset{\sim}{X}$.

2.4 The Asymptotic Null Distribution of T_n

The asymptotic normality of T_n under H_0 is a direct result of a theorem by Randles (1982), which gives the asymptotic normality of a U-statistic which involves an estimated parameter. To verify the conditions (given below) of Randles' theorem, we need the following assumptions:

2.4.1 E (E') is a continuous random variable with a bounded and continuous density function, has median zero and a finite variance.

2.4.2 The covariate term, X, has a finite fourth moment.

Consider the U-statistic

$$T_n(\underset{\sim}{\gamma}) = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{[(E_i-E_j)-(\gamma_1-\beta_1)(X_i-X_j)][(E_i'-E_j')-(\gamma_2-\beta_2)(X_i-X_j)]\} \tag{2.4.3}$$

where the mathematical variable $\underset{\sim}{\gamma} = (\gamma_1, \gamma_2)'$ replaces the estimator $\hat{\underset{\sim}{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$. The corresponding kernel is

$$h(\underset{\sim}{S}_1, \underset{\sim}{S}_2; \underset{\sim}{\gamma}) = \text{Sgn}\{[(E_1-E_2)-(\gamma_1-\beta_1)(X_1-X_2)][(E_1'-E_2')-(\gamma_2-\beta_2)(X_1-X_2)]\} \tag{2.4.4}$$

with $\underset{\sim}{S}_i = (X_i, E_i, E_i')'$. The Conditions of Randles' theorems are as follows:

Condition 2.4.5

$$n^{1/2}(\hat{\beta} - \beta) = O_p(1) .$$

Condition 2.4.6

Suppose there is a neighborhood of β, say $K(\beta)$, and a constant $K_1>0$ such that if $\gamma\varepsilon K(\beta)$ and $D(\gamma,d)$ is a sphere centered at γ with radius d satisfying $D(\gamma,d)\subset K(\beta)$, then

$$E[\sup_{\gamma'\varepsilon D(\gamma,d)} | h(S_1,S_2;\ \gamma') - h(S_1,S_2;\ \gamma) |] \leq K_1 d .$$

Condition 2.4.7

Suppose there exists a constant $M_1 > 0$ such that

$$|h(x_1,x_2;\ \gamma) - h(x_1,x_2;\ \beta)| \leq M_1$$

for every x_1, x_2 and for all γ in some neighborhood of β.

Condition 2.4.8

$\theta(\gamma)$ has a zero differential at $\gamma=\beta$, where

$$\theta(\gamma) = E[T_n(\gamma)] = E\ [h(S_1,S_2;\ \gamma)] .$$

Condition 2.4.9

$$n^{1/2}\ [T_n(\beta) - \theta(\beta)] \xrightarrow{d} N(0,\tau^2)$$

where $\tau^2 = 4\ \mathrm{Var}\{E[h(S_1,S_2;\ \beta)|S_1]\}$.

THEOREM 2.4.10

Under assumptions 2.4.1 and 2.4.2,

$$n^{1/2}[T_n(\hat{\beta}) - \Theta(\beta)] \xrightarrow{d} N(0,\tau^2) \text{ , as } n\to\infty \text{ .}$$

Proof. This is seen by verifying conditions 2.4.5-2.4.9 given above.

Condition 2.4.5

We need to show

$$n^{1/2}(\hat{\beta} - \beta) = O_p(1) \text{ .}$$

For this condition to hold in general one needs a stronger assumption than the consistency of the estimator $\hat{\beta}$.

For example, from the Markov inequality, and for i = 1, 2,

$$P\{ n^{1/2} \mid \hat{\beta}_i - \beta_i \mid > \varepsilon \} \leq nE[(\hat{\beta}_i - \beta_i)^2]/\varepsilon^2 \text{ ,}$$

so that it is sufficient that the second moment of $\hat{\beta}_i$ around β_i be of order $n^{-\delta}$, $\delta \geq 1$. However, in our particular setting, when $\hat{\beta}_i$ is the slope estimator of β_i in the simple linear model, we can show that for i = 1, 2, $n^{1/2}(\hat{\beta}_i - \beta_i)$ converges to some bona fide distribution, thereby proving this condition. In what follows we shall demonstrate that, under very broad assumptions, this indeed is the case for the two estimators of interest: (i) the OLS estimator and (ii) the LAV estimator $\hat{\beta}$ of β.

(i) The OLS estimator:

For the model

$$Y_i = \alpha + \beta X_i + E_i \quad , i = 1, 2, \ldots, n$$

$$\hat{\beta} - \beta = \frac{\sum_{i=1}^{n} (X_i - \bar{X})(E_i - \bar{E})}{\sum_{i=1}^{n} (X_i - \bar{X})^2} \; .$$

$$= \frac{\frac{1}{n} \sum_{i=1}^{n} X_i E_i - \bar{X}\,\bar{E}}{\frac{1}{n} \sum_{i=1}^{n} X_i^2 - \bar{X}^2}$$

$$\equiv \frac{C_4 - C_1 C_2}{C_3 - C_1^2}$$

$$= g(\underset{\sim}{C})$$

where

$$C_1 = \bar{X} \, , \quad C_2 = \bar{E} \, , \quad C_3 = \frac{1}{n} \sum_{i=1}^{n} X_i^2 \quad , \; C_4 = \frac{1}{n} \sum_{i=1}^{n} X_i E_i \; ,$$

and $\underset{\sim}{C}' = (C_1, C_2, C_3, C_4)$. Using properties of sample moments (see for example, Serfling, 1980, p. 125), we see that

$$\underset{\sim}{C} \text{ is } AN \, [E(\underset{\sim}{C}) \, , \frac{1}{n} \Sigma] \, ,$$

where Σ is the covariance matrix of $(X_1, E_1, X_1^2, X_1 E_1)$. By the independence of the X_i's and E_i's, $g(E(\underset{\sim}{C})) = 0$, so that by corollary

3.3 of Serfling

$$(\hat{\beta}-\beta) \equiv g(\underset{\sim}{C}) \text{ is AN } (0 , \frac{1}{n}\underset{\sim}{D}' \textstyle\sum \underset{\sim}{D}) ,$$

where

$$\underset{\sim}{D}' = (\frac{\partial g}{\partial C_1} \bigg|_{\underset{\sim}{C}=E(\underset{\sim}{C})} , \ldots, \frac{\partial g}{\partial C_4} \bigg|_{\underset{\sim}{C}=E(\underset{\sim}{C})}) .$$

Therefore,

$$n^{1/2}(\hat{\beta}-\beta) \overset{d}{\rightarrow} N (0, \underset{\sim}{D}' \textstyle\sum \underset{\sim}{D}), \text{ as } n \rightarrow \infty .$$

Hence, since X_1 admits a finite fourth moment, and the error terms have finite variance, $n^{1/2}(\hat{\beta}-\beta)$ converges in law to a bona fide distribution implying $n^{1/2}(\hat{\beta}-\beta) = O_p(1)$.

(ii) The LAV estimator:

Consider the linear model given in (i) above, and let $\underset{\sim}{\beta}^*$ be the LAV estimator of $\underset{\sim}{\beta} = (\alpha,\beta)'$, i.e., $\underset{\sim}{\beta}^* = (\alpha^*,\beta^*)'$ is a solution to

$$\min_{\underset{\sim}{\beta}\in R^2} \{ \sum_{i=1}^{n} | y_i-\alpha_i-\beta x_i |\} .$$

Let $H(\cdot)$ denote the absolutely continuous distribution function of E_i with median zero and continuous and positive density $h(\cdot)$ at zero. Let X_n denote the nx2 regression matrix which depends on n through the sequence of constants $x_1, x_2, \ldots , x_n$. Bassett and Koenker (1978) have shown that if $Q = \lim_{n\to\infty} \frac{1}{n} X_n'X_n$ is positive definite, then $n^{1/2}(\underset{\sim}{\beta}^*-\underset{\sim}{\beta})$ converges in distribution to a bivariate normal vector with mean $\underset{\sim}{0}$ and covariance matrix w^2Q^{-1}, where $w = [2h(0)]^{-1}$. The above result implies that for the slope estimator β^*, $n^{1/2}(\beta^*-\beta)$ converges in distribution to a normal random variable with mean zero and variance $\nu^2 = w^2\underset{\sim}{\lambda}'Q^{-1}\underset{\sim}{\lambda}$, where $\underset{\sim}{\lambda} = [0,1]'$. Letting

$$g_n(\underset{\sim}{x}_n,\underset{\sim}{E}_n) = n^{1/2}(\beta^*-\beta)/\nu$$

and

$$F_n(t) = P\{ g_n(\underset{\sim}{x}_n,E_n) \leq t \} ,$$

for every sequence of regression constants $\{x_n\}$ for which Q exists and is positive definite, we have

$$\lim_{n\to\infty} F_n(t) = \Phi(t)$$

where $\Phi(\cdot)$ denotes the distribution function of the standard normal random variable. In the case where $X_1, X_2, \ldots, X_n$ is a sequence of random variables defined on a probability space P and having mean zero and variance σ_X^2 ,

$$\frac{1}{n} X_n'X_n \overset{a.s.}{\to} Q = \begin{bmatrix} 1 & 0 \\ 0 & \sigma_X^2 \end{bmatrix} \quad \text{as } n \to \infty ,$$

where

$$\frac{1}{n} X_n'X_n = \begin{bmatrix} 1 & \bar{X}_n \\ \bar{X}_n & \sum_{i=1}^{n} \frac{X_i^2}{n} \end{bmatrix} ,$$

with

$$\bar{X}_n = \sum_{i=1}^{n} \frac{X_i}{n} \quad .$$

It follows that

$$Q^{-1} = \frac{1}{\sigma_X^2} \begin{bmatrix} \sigma_X^2 & 0 \\ 0 & 1 \end{bmatrix},$$

and

$$\nu = w^2 \underset{\sim}{\lambda}' Q^{-1} \underset{\sim}{\lambda} = w^2/\sigma_X^2 ,$$

so that if we let

$$F_{n,\underset{\sim}{X}_n}(t) = P \{ g_n(\underset{\sim}{X}_n, \underset{\sim}{E}_n) \leq t \mid \underset{\sim}{X}_n \} ,$$

by Basset and Koenker's (1978) result we have

$$\lim_{n\to\infty} F_{n,\underset{\sim}{X}_n}(t) = \Phi(t) \quad \text{a.e. in } \underset{\sim}{X}_n \ .$$

But

$$P \{ g_n(\underset{\sim}{X}_n, \underset{\sim}{E}_n) \leq t \}$$

$$= \int_{\underset{\sim}{X}_n} F_{n,\underset{\sim}{X}_n}(t) \, dP$$

$$= \int_{\underset{\sim}{X}_n} \{ I_{[g_n(\underset{\sim}{X}_n, \underset{\sim}{E}_n) \leq t]} \mid \underset{\sim}{X}_n \} \, dP$$

where $I_{[\cdot]}$ is an indicator function. Then by the Lebesgue Dominated Convergence theorem

$$\lim_{n\to\infty} P \{ g_n(\underset{\sim}{X}_n, \underset{\sim}{E}_n) \leq t \}$$

$$= \int_{\underset{\sim}{X}_n} \lim_{n\to\infty} F_{n,\underset{\sim}{X}_n}(t) \, dP = \Phi(t) ,$$

and therefore

$$n^{1/2}\sigma_X(\beta^*-\beta)/w \overset{d}{\to} N(0,1) \quad \text{as } n\to\infty.$$

Condition 2.4.6

To verify condition 2.4.6 for the kernel of this setting, we examine the following:

Let

$$S^* = \sup_{\underset{\sim}{\gamma}'\varepsilon D(\underset{\sim}{\gamma},d)} \left| h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\gamma}') - h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\gamma}) \right|$$

$$= \sup_{\underset{\sim}{\gamma}'\varepsilon D(\underset{\sim}{\gamma},d)} \{| \text{Sgn}\{[(E_1-E_2)-(\gamma_1'-\beta_1)(X_1-X_2)][(E_1'-E_2')-(\gamma_2'-\beta_2)(X_1-X_2)]\} - \text{Sgn}\{[(E_1-E_2)-(\gamma_1-\beta_1)(X_1-X_2)][(E_1'-E_2')-(\gamma_2-\beta_2)(X_1-X_2)]\}|\}.$$

Denoting $B_i(\underset{\sim}{\xi}) = (\xi_i-\beta_i)(X_1-X_2)$, i = 1, 2, we have

$$S^* = \begin{cases} 2 \text{ if} & [(E_1-E_2)-B_1(\underset{\sim}{\gamma}')][(E_1'-E_2')-B_2(\underset{\sim}{\gamma}')] > 0\ (<0) \\ \quad\text{and} & [(E_1-E_2)-B_1(\underset{\sim}{\gamma})][(E_1'-E_2')-B_2(\underset{\sim}{\gamma})] < 0\ (>0) \\ 1 \text{ if} & [(E_1-E_2)-B_1(\underset{\sim}{\gamma}')][(E_1'-E_2')-B_2(\underset{\sim}{\gamma}')] = 0\ (\neq 0) \\ \quad\text{and} & [(E_1-E_2)-B_1(\underset{\sim}{\gamma})][(E_1'-E_2')-B_2(\underset{\sim}{\gamma})] \neq 0\ (=0) \\ 0 & \text{otherwise}. \end{cases}$$

When taking expectations, only the value of S* = 2 contributes to the expected value, since for S* = 1 the expectation involves probabilities of continuous variables taking on zero values. Hence,

$$E[S^*] = 2P\{[(E_1-E_2)-B_1(\underset{\sim}{\gamma}')][(E_1'-E_2')-B_2(\underset{\sim}{\gamma}')] > 0 ,$$

$$[(E_1-E_2)-B_1(\underset{\sim}{\gamma})][(E_1'-E_2')-B_2(\underset{\sim}{\gamma})] < 0\}$$

$$+ 2P\{[(E_1-E_2)-B_1(\underset{\sim}{\gamma}')][(E_1'-E_2')-B_2(\underset{\sim}{\gamma}')] < 0 ,$$

$$[(E_1-E_2)-B_1(\underset{\sim}{\gamma})][(E_1'-E_2')-B_2(\underset{\sim}{\gamma})] > 0\}$$

$$= 2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') > 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') > 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) < 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) > 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') > 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') > 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) > 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) < 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') < 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') < 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) > 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) < 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') < 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') < 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) < 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) > 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') > 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') < 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) > 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) > 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') > 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') < 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) < 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) < 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') < 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') > 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) > 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) > 0\}$$

$$+2P\{(E_1-E_2)-B_1(\underset{\sim}{\gamma}') < 0, (E_1'-E_2')-B_2(\underset{\sim}{\gamma}') > 0, (E_1-E_2)-B_1(\underset{\sim}{\gamma}) < 0,$$

$$(E_1'-E_2')-B_2(\underset{\sim}{\gamma}) < 0\} \quad .$$

Denote the above probabilities by $p_1, p_2, p_3, \ldots, p_8$, so that $E[S^*] = 2(p_1+p_2+p_3+p_4+p_5+p_6+p_7+p_8)$, and note that for $X_1 > X_2$,

$$p_1 \leq P\ \{(E_1-E_2)-B_1(\underset{\sim}{\gamma}')>0,\ (E_1-E_2)-B_1(\underset{\sim}{\gamma})<0\ \}$$

$$= P\ \{(E_1-E_2)-(\gamma_1'-\beta_1)(X_1-X_2)>0,\ (E_1-E_2)-(\gamma_1-\beta_1)(X_1-X_2)<0\ \}$$

$$= P\ \{\ \frac{E_1-E_2}{X_1-X_2} > \gamma_1'-\beta_1,\ \ \frac{E_1-E_2}{X_1-X_2} < \gamma_1 - \beta_1\ \}\ .$$

Similarly,

$$p_3 \leq P\ \{\ \frac{E_1-E_2}{X_1-X_2} < \gamma_1' - \beta_1\ ,\ \ \frac{E_1-E_2}{X_1-X_2} > \gamma_1 - \beta_1\ \}\ \ .$$

By assumptions 2.4.1 and 2.4.2 the random variable $\frac{E_1-E_2}{X_1-X_2}$ $[\frac{E_1'-E_2'}{X_1-X_2}]$ has a distribution function $K(\cdot)[K'(\cdot)]$, and a density $k(\cdot)[k'(\cdot)]$ which is bounded by B[B'] and continuous, so that

$$P_1 + P_3 \leq 2P\{\min(\gamma_1' - \beta_1, \gamma_1 - \beta_1) < \frac{E_1 - E_2}{X_1 - X_2} < \max(\gamma_1' - \beta_1, \gamma_1 - \beta_1)\}$$

$$= 2K[\max(\gamma_1' - \beta_1, \gamma_1 - \beta_1)] - 2K[\min(\gamma_1' - \beta_1, \gamma_1 - \beta_1)]$$

$$= 2 \mid \max(\gamma_1' - \beta_1, \gamma_1 - \beta_1) - \min(\gamma_1' - \beta_1, \gamma_1 - \beta_1) \mid k(\xi^*)$$

$$= 2 \mid \max(\gamma_1', \gamma_1) - \min(\gamma_1', \gamma_1) \mid k(\xi^*)$$

$$= 2\, d\, B ,$$

where $\xi^* = \delta[\max(\gamma_1', \gamma_1) - \min(\gamma_1', \gamma_1)]$ for $|\delta| < 1$, and since $\underset{\sim}{\gamma}' \in D(\underset{\sim}{\gamma}, d)$.

Similarly $P_2 + P_4 \leq 2dB'$, $P_5 + P_8 \leq 2dB'$, and $P_6 + P_7 \leq 2dB$, so that $E[S^*] \leq 8d(B+B')$, which proves condition 2.4.6 with $K_1 = 8(B+B')$.

Condition 2.4.7

By the definition of the kernel h, this condition holds with $M_1 = 2$.

Condition 2.4.8

We need to show that

$$\Theta(\underset{\sim}{\gamma}) \text{ has a zero differential at } \underset{\sim}{\gamma} = \underset{\sim}{\beta},$$

where $\Theta(\underset{\sim}{\gamma}) = E[T_n(\underset{\sim}{\gamma})] = E[h(\underset{\sim}{S}_1, \underset{\sim}{S}_2; \underset{\sim}{\gamma})]$.

Using the notation adopted under condition 2.4.6, and conditioning on X_1 and X_2,

$$\theta(\underset{\sim}{\gamma}) = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\gamma})|X_1 = x_1,\ X_2 = x_2]$$

$$= E[\mathrm{Sgn}\{[(E_1-E_2)-b_1(\underset{\sim}{\gamma})][(E_1'-E_2')-b_2(\underset{\sim}{\gamma})]\}]$$

$$= P\ \{\ (E_1-E_2) > b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') > b_2(\underset{\sim}{\gamma})\ \}$$

$$+P\ \{\ (E_1-E_2) < b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') < b_2(\underset{\sim}{\gamma})\ \}$$

$$-P\ \{\ (E_1-E_2) > b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') < b_2(\underset{\sim}{\gamma})\ \}$$

$$-P\ \{\ (E_1-E_2) < b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') > b_2(\underset{\sim}{\gamma})\ \}$$

$$= 2P\ \{\ (E_1-E_2) > b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') > b_2(\underset{\sim}{\gamma})\ \}$$

$$+2P\ \{\ (E_1-E_2) < b_1(\underset{\sim}{\gamma}),\ (E_1'-E_2') < b_2(\underset{\sim}{\gamma})\ \} - 1\ .$$

Under the null hypothesis of the independence of E_i and E_i', i = 1, 2, . . . , n, the above probabilities factor to yield

$$E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\gamma})|X_1 = x_1,\ X_2 = x_2]$$

$$= 2[1-F_1(b_1(\underset{\sim}{\gamma}))][1-F_2(b_2(\underset{\sim}{\gamma}))] + 2F_1(b_1(\underset{\sim}{\gamma}))F_2(b_2(\underset{\sim}{\gamma})) - 1].$$

Thus

$$\theta(\underset{\sim}{\gamma}) = E_{X_1,X_2}\{2[1-F_1(B_1(\underset{\sim}{\gamma}))][1-F_2(B_2(\underset{\sim}{\gamma}))] + 2F_1(B_1(\underset{\sim}{\gamma}))F_2(B_2(\underset{\sim}{\gamma})) - 1\}$$

where

$$B_i(\underset{\sim}{\xi}) = (\xi_i - \beta_i)(X_1 - X_2) \ , \ i = 1,\ 2,$$

and E_{X_1,X_2} denotes a two-fold integral yielding the expectation with respect to X_1, X_2. Using assumptions 2.4.1 and 2.4.2, differentiation with respect to $\underset{\sim}{\gamma}$ may be passed inside the integral (see for example theorem A.2.4 of Randles and Wolfe, 1979), yielding the differential of the function $\Theta(\underset{\sim}{\gamma})$ to be

$$\begin{aligned}
\partial\Theta(\underset{\sim}{\gamma}) = E_{X_1,X_2}\{ & -2\partial B_1(\underset{\sim}{\gamma}) f_1(B_1(\underset{\sim}{\gamma}))[1-F_2(B_2(\underset{\sim}{\gamma}))] \\
& -2\partial B_2(\underset{\sim}{\gamma}) f_2(B_2(\underset{\sim}{\gamma}))[1-F_1(B_1(\underset{\sim}{\gamma}))] \\
& +2\partial B_1(\underset{\sim}{\gamma}) f_1(B_1(\underset{\sim}{\gamma}))F_2(B_2(\underset{\sim}{\gamma})) \\
& +2\partial B_2(\underset{\sim}{\gamma}) f_2(B_2(\underset{\sim}{\gamma}))F_1(B_1(\underset{\sim}{\gamma}))\} \\
= E_{X_1,X_2}\{ & 2\partial B_1(\underset{\sim}{\gamma}) f_1(B_1(\underset{\sim}{\gamma}))[2F_2(B_2(\underset{\sim}{\gamma}))-1] \\
& +2\partial B_2(\underset{\sim}{\gamma}) f_2(B_2(\underset{\sim}{\gamma}))[2F_1(B_1(\underset{\sim}{\gamma}))-1]\} \\
= 0 \quad & \text{at } \underset{\sim}{\gamma} = \underset{\sim}{\beta} \ , \text{ since } B_i(\underset{\sim}{\beta}) = 0,
\end{aligned}$$

$i = 1,\ 2$, and $F_1(0) = F_2(0) = 1/2$. This proves condition 2.4.8.

Condition 2.4.9

We need to show that

$$n^{1/2}[T_n(\underset{\sim}{\beta}) - \Theta(\underset{\sim}{\beta})] \overset{d}{\rightarrow} N\ (0, \tau^2) \ .$$

This is a direct consequence of U-statistics theorems, since under H_0

$$T_n(\underset{\sim}{\beta}) = \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{(E_i - E_j)(E_i' - E_j')\}$$

is a U-statistic based on the i.i.d. random variables $(E_1,E_1'), \ldots,$ (E_n,E_n'). (See for example Theorem 3.3.13, Randles and Wolfe, 1979.) Further, under H_0,

$$\theta(\underset{\sim}{\beta}) = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})]$$

$$= 2P\{E_1-E_2>0,\ E_1'-E_2'>0\} + 2P\{E_1-E_2<0,\ E_1'-E_2'<0\} - 1$$

$$= 2P\{E_1-E_2>0\}\cdot P\{E_1'-E_2'>0\} + 2P\{E_1-E_2<0\}\cdot P\{E_1'-E_2'<0\} - 1$$

$$= 0\ ,$$

$$\tau^2 = 4\mathrm{Var}\ \{E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})|\underset{\sim}{S}_1 = \underset{\sim}{s}\ \}$$

$$= 4\ E_{\underset{\sim}{S}_1}\ \{(E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})|\underset{\sim}{S}_1])^2\}\ ,$$

since

$$E_{\underset{\sim}{S}_1}\ \{E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})|\underset{\sim}{S}_1]\} = E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})] = 0\ .$$

With $\underset{\sim}{S}_1 = \underset{\sim}{s} \equiv (t,u,v)$, and by procedures similar to those given in verifying condition 2.4.8,

$$E[h(\underset{\sim}{S}_1,\underset{\sim}{S}_2;\ \underset{\sim}{\beta})|\underset{\sim}{S}_1 = \underset{\sim}{s}] = 2P\{(u-E_2) > 0,\ (v-E_2') > 0\}$$

$$+ 2P\{(u-E_2) < 0,\ (v-E_2') < 0\} - 1$$

$$= 2H_1(u)H_2(v) + 2[1-H_1(u)][1-H_2(v)] - 1,$$

where $H_1(\cdot)$ ($H_2(\cdot)$) is the distribution function of $E(E')$. Therefore,

$$\tau^2 = 4 \int_0^1\int_0^1 \{ 4H_1^2(u)H_2^2(v) + 4[1-H_1(u)]^2[1-H_2(v)]^2$$

$$+ 1 + 8\, H_1(u)H_2(v)[1-H_1(u)][1-H_2(v)] - 4H_1(u)H_2(v)$$

$$- 4[1-H_1(u)][1-H_2(v)]\}\ dH_1(u)dH_2(v)\ .$$

The above expression contains four types of terms:

(i) $$\int_0^1\int_0^1 H_1^2(u)H_2^2(v)dH_1(u)dH_2(v) = \frac{1}{9}\, H_1^3(u)\Big|_0^1 . H_2^3(v)\Big|_0^1 = \frac{1}{9}$$

(ii) $$\int_0^1\int_0^1 [1-H_1(u)]^2[1-H_2(v)]^2 dH_1(u)dH_2(v) =$$

$$= \frac{1}{9}\, [1-H_1(u)]^3 \Big|_0^1\ [1-H_2(v)]^3 \Big|_0^1 = \frac{1}{9}$$

(iii) $$\int_0^1\int_0^1 [H_1(u)-H_1^2(u)][H_2(v)-H_2^2(v)]dH_1(u)dH_2(v)$$

$$= [\ \frac{1}{2}\, H_1^2(u) - \frac{1}{2}\, H_1^3(u)]\ \Big|_0^1\ [\ \frac{1}{2}\, H_2^2(v) - \frac{1}{3}\, H_2^3(v)]\ \Big|_0^1 = \frac{1}{36}\ ,$$

and

(iv) $\int_0^1 \int_0^1 H_1(u)H_2(v)dH_1(u)dH_2(v) = \frac{1}{4} H_1^2(u)H_2^2(v) \Big|_0^1 = \frac{1}{4}$

$$= \int_0^1 \int_0^1 [1-H_1(u)][1-H_2(v)]dH_1(u)dH_2(v) .$$

Therefore,

$$\tau^2 = 4[4(\frac{1}{9}) + 4(\frac{1}{9}) + 1 + \frac{8}{36} - 4(\frac{1}{4}) - 4(\frac{1}{4})] = \frac{4}{9} .$$

Conditions 2.4.5-2.4.9 are satisfied so that by Randles' theorem (1982), under H_0

$$n^{1/2}\, T_n(\hat{\underset{\sim}{\beta}}) \overset{d}{\rightarrow} N\,(0, \frac{4}{9}) , \quad \text{as } n\rightarrow\infty .$$

2.5 The Simulated Null Distribution of T_n Under Normality

The tables in this section contain the empirical null distributions of T_n obtained by a Monte Carlo simulation study. This and all other studies in subsequent chapters were performed on the University of Florida IBM-3033 using Fortran. A copy of some main programs and subroutines used in this work is given in the appendix.

In generating the distribution of T_n under the hypothesis of the independence of E and E', we bear in mind that the distribution of T_n is free of the regression constants involved in the underlying linear models (2.1.1), and of the location and scale parameters of X, E and

E', as discussed in section 2.2. The simulated distribution of T_n is then obtained as follows: the IMSL subroutine GGNML is used to generate 3n i.i.d. random variables from the standard normal distribution. These are then divided into three groups of size n each to yield X_i, E_i and E_i', i = 1, 2, . . . , n, and the following models are obtained

$$Y_i = X_i + E_i$$

and

$$Z_i = X_i + E_i' \quad , \; i=1,2,\ldots,n \; .$$

From these models we obtain residual pairs in two ways: (i) by the ordinary least squares (OLS) procedures, and (ii) by the least absolute value (LAV) method. The LAV estimates of the regression parameters were obtained by an algorithm given by Josvanger and Sposito (1983). This algorithm is reproduced in the appendix. In each of the two cases (the OLS and the LAV), the usual Kendall's tau was calculated on the residuals. This process was repeated 10,000 times, and the frequency distributions for the different possible values of the statistic were recorded. The empirical relative frequency distributions of T_n for the two cases are given in Tables 2.1 and 2.2, respectively.

Table 2.1

The Null Distribution of T_n (OLS fit)

For a given n, the entry in the table for the point x is $\hat{\alpha}$, the empirical estimate of $\alpha=P_o[\binom{n}{2}T_n \geq x]$, where T_n is obtained from the residuals of an OLS fit.

	n							
x	6	7	10	11	14	15	18	19
1	.4991	.5000	.5003	.5012	.4998	.4890	.5043	.5034
3	.3773	.3992	.4361	.4438	.4603	.4569	.4739	.4752
5	.2667	.3049	.3715	.3908	.4197	.4229	.4422	.4497
7	.1702	.2174	.3100	.3364	.3803	.3871	.4143	.4250
9	.0968	.1451	.2537	.2854	.3428	.3504	.3874	.3979
11	.0478	.0877	.2022	.2378	.3073	.3157	.3581	.3719
13	.0187	.0492	.1628	.1927	.2721	.2831	.3311	.3460
15	.0040	.0237	.1221	.1544	.2400	.2534	.3066	.3173
17		.0103	.0901	.1228	.2055	.2249	.2810	.2918
19		.0033	.0631	.0956	.1768	.1978	.2588	.2725
21		.0006	.0444	.0746	.1493	.1697	.2367	.2533
23			.0294	.0544	.1264	.1477	.2164	.2322
25			.0195	.0391	.1055	.1264	.1958	.2128
27			.0128	.0265	.0889	.1086	.1757	.1925
29			.0083	.0161	.0713	.0908	.1572	.1738
31			.0041	.0110	.0578	.0749	.1375	.1563
33			.0023	.0066	.0456	.0611	.1233	.1401
35			.0012	.0047	.0362	.0510	.1101	.1257
37				.0028	.0287	.0409	.0981	.1123
39				.0012	.0229	.0336	.0874	.0994
41				.0007	.0158	.0251	.0756	.0863
43				.0004	.0125	.0202	.0665	.0754
45					.0086	.0160	.0565	.0657
47					.0056	.0122	.0476	.0579
49					.0045	.0087	.0404	.0500
51					.0031	.0071	.0326	.0423
53					.0020	.0053	.0267	.0363
55					.0009	.0038	.0226	.0309
57					.0005	.0026	.0182	.0272
59					.0003	.0019	.0152	.0231
61						.0017	.0121	.0193
63						.0012	.0092	.0167

Table 2.1-continued.

x	n = 6	7	10	11	14	15	18	19
65						.0008	.0071	.0135
67							.0052	.0109
69							.0042	.0093
71							.0029	.0082
73							.0021	.0064
75							.0016	.0054
77							.0013	.0044
79							.0011	.0040
81							.0010	.0029
83							.0008	.0023
85							.0007	.0021
87								.0018
89								.0011
91								.0008
93								.0007

x	n = 4	5	8	9	12	13	16	17	20
0	.5867	.5785	.5482	.5434	.5329	.5236	.5182	.5114	.5152
2	.4089	.4299	.4620	.4648	.4847	.4787	.4829	.4812	.4932
4	.2503	.2809	.3703	.3922	.4341	.4296	.4492	.4493	.4677
6	.1056	.1630	.2907	.3190	.3827	.3855	.4134	.4185	.4404
8	.0000	.0751	.2241	.2550	.3346	.3441	.3833	.3897	.4203
10		.0229	.1596	.1965	.2865	.3022	.3494	.3570	.3954
12			.1091	.1460	.2437	.2667	.3143	.3280	.3682
14			.0701	.1063	.2069	.2296	.2844	.2961	.3463
16			.0420	.0717	.1742	.1945	.2568	.2724	.3253
18			.0239	.0468	.1413	.1616	.2298	.2482	.3020
20			.123	.0275	.1138	.1339	.2067	.2241	.2823
22			.0045	.0157	.0884	.1070	.1823	.1999	.2606
24			.0018	.0095	.0652	.0870	.1582	.1787	.2388
26			.0008	.0050	.0504	.0709	.1380	.1597	.2218
28				.0022	.0395	.0543	.1195	.1419	.2049
30				.0007	.0293	.0424	.1025	.1244	.1868
32					.0202	.0337	.0873	.1083	.1697
34					.0142	.0260	.0759	.0949	.1537
36					.0097	.0194	.0632	.0803	.1396

Table 2.1-continued.

	n								
x	4	5	8	9	12	13	16	17	20
38					.0072	.0136	.0518	.0691	.1253
40					.0048	.0089	.0434	.0599	.1117
42					.0027	.0056	.0364	.0511	.1004
44					.0014	.0044	.0305	.0423	.0907
46					.0009	.0028	.0240	.0347	.0789
48					.0003	.0018	.0202	.0276	.0701
50						.0012	.0162	.0226	.0625
52						.0006	.0134	.0182	.0546
54							.0103	.0152	.0469
56							.0076	.0120	.0416
58							.0060	.0094	.0371
60							.0048	.0076	.0326
62							.0039	.0059	.0283
64							.0033	.0043	.0241
66							.0019	.0033	.0205
68							.0016	.0025	.0169
70							.0014	.0023	.0141
72							.0010	.0016	.0120
74								.0011	.0098
76								.0006	.0086
78								.0004	.0067
80									.0052
82									.0038
84									.0033
86									.0023
88									.0019
90									.0017
92									.0016
94									.0011
96									.0010
98									.0009

Table 2.2

The Null Distribution of T_n (LAV fit)

For a given n, the entry in the table for the point x is $\hat{\alpha}$, the empirical estimate of $\alpha = P_0[\binom{n}{2} T_n \geq x]$, where T_n is obtained from the residuals of an LAV fit.

	n							
x	6	7	10	11	14	15	18	19
1	.5208	.5038	.5072	.5097	.5041	.5072	.5044	.5015
3	.3766	.3930	.4349	.4458	.4641	.4672	.4741	.4742
5	.2499	.2958	.3724	.3903	.4193	.4299	.4439	.4467
7	.1528	.2081	.3103	.3338	.3735	.3936	.4166	.4212
9	.0938	.1381	.2506	.2836	.3323	.3579	.3888	.3955
11	.0559	.0844	.2010	.2323	.2957	.3214	.3589	.3692
13	.0263	.0493	.1554	.1895	.2633	.2890	.3344	.3460
15	.0057	.0259	.1163	.1547	.2298	.2595	.3073	.3214
17	.0000	.0121	.0878	.1176	.1988	.2296	.2816	.2966
19		.0037	.0612	.0888	.1693	.2021	.2562	.2753
21		.0009	.0414	.0642	.1441	.1757	.2332	.2522
23		.0000	.0277	.0465	.1201	.1541	.2120	.2304
25			.0182	.0323	.1002	.1308	.1921	.2099
27			.0110	.0230	.0829	.1086	.1707	.1896
29			.0061	.0159	.0685	.0892	.1510	.1695
31			.0029	.0095	.0548	.0759	.1357	.1540
33			.0017	.0052	.0434	.0617	.1224	.1367
35			.0009	.0036	.0339	.0501	.1082	.1200
37			.0005	.0020	.0246	.0416	.0944	.1069
39			.0004	.0014	.0180	.0326	.0810	.0955
41				.0005	.0128	.0260	.0697	.0829
43				.0004	.0093	.0197	.0593	.0734
45				.0002	.0065	.0162	.0496	.0648
47				.0000	.0042	.0127	.0413	.0564
49					.0028	.0093	.0351	.0495
51					.0022	.0070	.0284	.0420
53					.0017	.0055	.0242	.0353
55					.0011	.0036	.0194	.0289
57					.0008	.0030	.0154	.0246
59					.0005	.0021	.0129	.0201
61					.0004	.0017	.0094	.0172
63					.0003	.0011	.0077	.0142

Table 2.2-continued.

x	n: 6	7	10	11	14	15	18	19
65						.0007	.0059	.0122
67						.0001	.0046	.0107
69						.0001	.0034	.0091
71							.0024	.0076
73							.0019	.0062
75							.0015	.0053
77							.0012	.0045
79							.0008	.0035
81							.0007	.0029
83							.0007	.0019
85							.0003	.0015
87							.0001	.0010
89							.0000	.0008
91								.0006
93								.0004
95								.0003
97								.0003
99								.0003

x	n: 4	5	8	9	12	13	16	17	20
0	.5652	.5887	.5417	.5362	.5267	.5245	.5089	.5157	.5170
2	.3976	.4258	.4511	.4620	.4776	.4769	.4768	.4829	.4935
4	.2805	.2801	.3604	.3911	.4236	.4329	.4416	.4509	.4674
6	.1269	.1617	.2823	.3161	.3754	.3868	.4117	.4236	.4407
8	.0000	.0882	.2106	.2493	.3211	.3394	.3758	.3909	.4161
10		.0312	.1440	.1928	.2746	.2991	.3459	.3588	.3916
12		.0000	.0971	.1421	.2372	.2600	.3134	.3285	.3687
14			.0607	.1037	.1984	.2246	.2827	.2990	.3455
16			.0360	.0700	.1628	.1916	.2555	.2708	.3230
18			.0211	.0455	.1339	.1599	.2276	.2458	.2997
20			.0112	.0295	.1106	.1333	.2029	.2219	.2795
22			.0061	.0185	.0893	.1060	.1812	.1995	.2618
24			.0029	.0099	.0686	.0863	.1554	.1798	.2427
26			.0007	.0052	.0514	.0691	.1354	.1590	.2225
28			.0000	.0031	.0373	.0539	.1173	.1418	.2020
30				.0016	.0259	.0421	.1001	.1246	.1832

Table 2.2-continued.

	n								
x	4	5	8	9	12	13	16	17	20
32				.0008	.0182	.0327	.0855	.1097	.1671
34				.0003	.0129	.0236	.0702	.0967	.1490
36					.0083	.0177	.0581	.0838	.1365
38					.0055	.0139	.0489	.0737	.1241
40					.0034	.0103	.0403	.0631	.1114
42					.0018	.0073	.0332	.0543	.1002
44					.0013	.0051	.0270	.0453	.0869
46					.0006	.0036	.0206	.0385	.0760
48					.0003	.0023	.0169	.0312	.0674
50					.0001	.0017	.0129	.0256	.0589
52						.0010	.0104	.0217	.0516
54						.0005	.0083	.0167	.0437
56						.0004	.0061	.0126	.0387
58						.0003	.0043	.0093	.0349
60							.0038	.0070	.0308
62							.0026	.0052	.0278
64							.0024	.0039	.0245
66							.0017	.0028	.0212
68							.0014	.0023	.0173
70							.0010	.0019	.0144
72							.0005	.0013	.0113
74							.0003	.0012	.0092
76								.0007	.0073
78								.0006	.0059
80								.0004	.0046
82								.0003	.0035
84									.0029
86									.0021
88									.0021
90									.0017
92									.0015
94									.0014
96									.0011
98									.0010
100									.0010
102									.0009
104									.0005
106									.0005
108									.0002

CHAPTER THREE
THE ASYMPTOTIC EFFICIENCY OF T_n RELATIVE TO THE PEARSON PARTIAL CORRELATION COEFFICIENT

3.1 Introduction

When investigating the performance of statistical tests for independence, the researcher is faced with the crucial problem of specifying an appropriate class of alternatives which is (i) sufficiently wide to encompass a large variety of situations, and (ii) is mathematically manageable. In our setting, this problem is further complicated by the presence of the slope estimators which induce dependence among the residual pairs (U_i, V_i), $i = 1, 2, \ldots, n$. To attain maximum generality and at the same time keep our investigation mathematically manageable, we adopt the "trivariate reduction" model for the errors. This is the model recommended by Hájek and Šidák (1967) for parametrizing the class of alternatives to the hypothesis of independence. Similar models were also considered by Konijn (1956) and Shirahata (1977).

The class of alternatives is constructed as follows:

let $$E_i = W_{1i} + \Delta W_{3i}$$

and

$$E_i' = W_{2i} + \Delta W_{3i}\ ,\ i=1,2,\ldots,n.$$

where $\{W_{1i}\}$, $\{W_{2i}\}$ and $\{W_{3i}\}$, $i = 1, 2, \ldots, n$ are three independent random samples of continuous random variables. The hypothesis that E_i and E_i' are independent is equivalent to the hypothesis that $\Delta = 0$, so that the test is equivalently given by

$$H_0 : \Delta = 0 \qquad \text{versus} \qquad H_a : \Delta \neq 0 .$$

To study the Pitman asymptotic relative efficiency (ARE), we will further suppose that Δ_n is a sequence of parameters converging to the null hypothesis value, i.e., $\lim_{n\to\infty} \Delta_n = 0$.

In section 3.2, we give a main result which ensures the asymptotic normality of a U-statistic with an estimated parameter under a sequence of alternatives converging to the null hypothesis. In section 3.3, we shall apply the results of section 3.2 to obtain the asymptotic normality of T_n, and in section 3.4 we derive the asymptotic distribution of the partial correlation coefficient, R_n. Section 3.5 contains the applications of a theorem by Noether, by which an expression for the asymptotic efficiency of T_n relative to R_n is obtained. A table of ARE's calculated for several underlying distributions is given at the end of section 3.5.

3.2 The Asymptotic Normality of a U-statistic with an Estimated Parameter Under a Sequence of Alternatives

The main result in this section is an extension of a theorem by Randles (1982) which involves a generalization of a result given by Sukhatme (1958). Randles' theorem is slightly modified to apply to the more general case where the U-statistic, U_n, and its moments are functions of the sample size, n, through the observations $X_{1:n}$, $X_{2:n}$, . . . , $X_{n:n}$, whose distribution in turn depends on n perhaps through a sequence of parameters Δ_n.

Let $X_{1:n}$, $X_{2:n}$, . . . , $X_{n:n}$ denote a random sample from some distribution with distribution function $F_n(x)$, possibly changing as n changes, and let $h(x_1, \ldots, x_r; \underset{\sim}{\gamma})$ denote a symmetric kernel of degree r with expected value

$$\theta_n(\underset{\sim}{\gamma}) = E_{\underset{\sim}{\beta}}[h(X_{1:n},\ldots, X_{r:n}; \underset{\sim}{\gamma})] ,$$

where $\underset{\sim}{\beta}$ denotes a P-dimensional parameter value, and $\underset{\sim}{\gamma}$ is, in general, a mathematical variable. Both the kernel and its expected value may depend on $\underset{\sim}{\gamma}$, and on n through $X_{1:n}$, . . . , $X_{n:n}$. The corresponding U-statistic is then

$$U_n(\underset{\sim}{\gamma}) = \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\epsilon A} h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n}; \underset{\sim}{\gamma}) , \qquad (3.2.1)$$

where A denotes the collection of all subsets of size r from the set of integers {1, 2, . . . , n}. The main result of this section gives the asymptotic normality of

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] ,$$

where $\hat{\underset{\sim}{\beta}}$ is an estimator of the parameter $\underset{\sim}{\beta}$. The key step in the proof of the main result requires that

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\hat{\underset{\sim}{\beta}}) - U_n(\underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})] \overset{P}{\to} 0 \text{ , as } n\to\infty \text{ ,} \qquad (3.2.2)$$

The proof of (3.2.2) is given in theorem 3.2.8, but first we prove a lemma and list the conditions needed for the proof of 3.2.2.

Lemma 3.2.3

Let $\underset{\sim}{X}_{1:n}, \underset{\sim}{X}_{2:n}, \ldots, \underset{\sim}{X}_{n:n}$ be i.i.d. random variables whose distribution may depend on n. Suppose $\tilde{k}_n(\underset{\sim}{X}_{1:n}, \ldots, \underset{\sim}{X}_{r:n})$ satisifies

(i) $E[\tilde{k}_n(\underset{\sim}{X}_{1:n},\ldots, \underset{\sim}{X}_{r:n})] = 0$, for every n , and

(ii) $\lim\limits_{n\to\infty} \frac{1}{n} E[\{\tilde{k}_n(\underset{\sim}{X}_{1:n},\ldots, \underset{\sim}{X}_{r:n})\}^2] = 0$,

then

$$U_n = \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} \tilde{k}_n(\underset{\sim}{X}_{\alpha_1:n},\ldots, \underset{\sim}{X}_{\alpha_r:n}) \overset{P}{\to} 0, \text{ as } n\to\infty$$

where A is as defined earlier.

Proof. Write

$$E[U_n^2] = \text{Var}[U_n] = \frac{1}{\binom{n}{r}} \sum_{c=1}^{r} \binom{r}{c}\binom{n-r}{r-c} \zeta_{c,n} \text{ ,}$$

where

$$\zeta_{c,n} = E[\tilde{k}_n(\underset{\sim}{X}_{1:n},\ldots,\underset{\sim}{X}_{r:n})\tilde{k}_n(\underset{\sim}{X}_{1:n},\ldots,\underset{\sim}{X}_{c:n},\underset{\sim}{X}_{r+1:n},\ldots,\underset{\sim}{X}_{2r-c:n})],$$

(see, for example, Randles and Wolfe, 1979, p. 65). Also, it has been shown that, for fixed n,

$$\zeta_{c,n} \leq \zeta_{r,n} \qquad \text{for } c = 1,2,\ldots,r \ ,$$

with

$$\zeta_{r,n} = E[\{\tilde{k}_n(\underset{\sim}{X}_{1:n},\ldots,\underset{\sim}{X}_{r:n})\}^2] \ .$$

Now, define K_c by

$$K_c = \frac{(r!)^2}{c![(r-c)!]^2} \ ,$$

so that each term in the above sum involves

$$K_c \cdot \frac{(n-r)(n-r-1)\ \ldots\ (n-2r+c+1)}{n(n-1)\ \ldots\ (n-r+1)}\ \zeta_{c,n}$$

$$\leq\ K_c \cdot \frac{(n-r)(n-r-1)\ \ldots\ (n-2r+c+1)}{n(n-1)\ \ldots\ (n-r+1)}\ \zeta_{r,n} \ .$$

Note that the numerator involves (r-c) factors of n, whereas the denominator involves r such factors, so that for each c = 1, 2, . . . , r, the coefficient of $\zeta_{r,n}$ is $O(n^{-\delta})$, $\delta \geq 1$, and therefore, from (ii), each term in the sum goes to zero, as n goes to infinity. It follows that, as n approaches infinity, $E[U_n^2]$ goes to zero, and, therefore, U_n converges in probability to zero.

Condition 3.2.4 Suppose

$$n^{1/2}(\hat{\beta} - \beta) = O_p(1) \quad \text{as } n\to\infty .$$

Condition 3.2.5 Suppose there is a neighborhood of β, say $K(\beta)$, and a positive constant K_1 such that if $\gamma \varepsilon K(\beta)$ and $D(\gamma,d)$ is a sphere centered at γ with radius d satisfying $D(\gamma,d) \subset K(\beta)$, then, for every n,

$$E[\sup_{\gamma' \varepsilon D(\gamma,d)} |h(X_{1:n},\ldots,X_{r:n};\, \gamma')-h(X_{1:n},\ldots,X_{r:n};\, \gamma)|] \leq K_1 d \quad (3.2.6)$$

and

$$\lim_{d\to 0} E[\sup_{\gamma' \varepsilon D(\gamma,d)} |h(X_{1:n},\ldots,X_{r:n};\, \gamma')-h(X_{1:n},\ldots,X_{r:n};\, \gamma)|^2] = 0$$

(3.2.7)

uniformly in n. That is, for every $\varepsilon'>0$ and every n, there exists a positive constant D' such that for $0<d<D'$ and $D(\gamma,d) \quad K(\beta)$,

$$E[\sup_{\gamma' \varepsilon D(\gamma,d)} |h(X_{1:n},\ldots,X_{r:n};\, \gamma')-h(X_{1:n},\ldots,X_{r:n};\, \gamma)|^2] < \varepsilon' .$$

THEOREM 3.2.8

Under conditions 3.2.4 and 3.2.5,

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\hat{\underset{\sim}{\beta}}) - U_n(\underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})] \overset{P}{\to} 0 \text{ , as } n\to\infty \text{ .}$$

PROOF. Let

$$\tilde{h}_n(x_1,\ldots, x_r; \underset{\sim}{\gamma}) = h(x_1,\ldots, x_r; \underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\gamma}) \text{ ,}$$

so that

$$U_n(\underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\gamma}) = \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots, X_{\alpha_r:n}; \underset{\sim}{\gamma})] \text{ ,}$$

where A denotes the collection of all subsets of r integers from $\{1, 2, \ldots, n\}$. Then,

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\hat{\underset{\sim}{\beta}}) - U_n(\underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})]$$

$$= \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n}; \hat{\underset{\sim}{\beta}}) - \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n}; \underset{\sim}{\beta})]$$

$$= \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n}; \underset{\sim}{\beta}+n^{-1/2}\hat{\underset{\sim}{s}}) - \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n}; \underset{\sim}{\beta})] \text{ .}$$

Denote the above expression by $Q_n(\hat{\underset{\sim}{s}})$, where $\hat{\underset{\sim}{s}} = n^{1/2}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})$. Now take ε and δ to be arbitrary constants. By Condition 3.2.4, $n^{1/2}(\hat{\underset{\sim}{\beta}}-\underset{\sim}{\beta})=O_p(1)$ so that we can find a sphere C in R^p centered at the origin, such that

$$P[n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \notin C] \leq \frac{\delta}{2}, \text{ for every } n . \tag{3.2.9}$$

Then,

$$P[|Q_n(n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}))| > \varepsilon]$$

$$= P[|Q_n(n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}))| > \varepsilon , n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \in C]$$

$$+ P[|Q_n(n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}))| > \varepsilon , n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \notin C]$$

$$\leq P[|Q_n(n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}))| > \varepsilon , n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \in C] + P[n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \notin C]$$

$$\leq P[\sup_{\underset{\sim}{s} \in C} | Q_n(\underset{\sim}{s}) | > \varepsilon] + P[n^{1/2}(\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) \notin C] .$$

It suffices to show that

$$P[\sup_{\underset{\sim}{s} \in C} | Q_n(\underset{\sim}{s}) | > \varepsilon] \to 0 \quad \text{as } n\to\infty ,$$

where ε and C are fixed.

Let C_u, u = 1, 2, . . . , U denote a finite collection of open spheres centered at $\underset{\sim}{s}_u$ with radii $\|C_u\| \leq \frac{\varepsilon}{8K_1}$ for every

$u = 1, 2, \ldots, U$, such that $\bigcup_u C_u \supset C$. Now,

$$Q_n(\underset{\sim}{s}) = \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta})]$$

$$= \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)]$$

$$+ \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)-\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta})]$$

$$\equiv Q_{n,u}(\underset{\sim}{s}) + Q_{n,o}(\underset{\sim}{s}_u) .$$

Also note that

$$P[\sup_{\underset{\sim}{s}\varepsilon C} | Q_n(\underset{\sim}{s}) | > \varepsilon] \leq \sum_{u=1}^{U} P[\sup_{\underset{\sim}{s}\varepsilon C_u} | Q_n(\underset{\sim}{s}) | > \varepsilon] ,$$

since

$$\{\sup_{\underset{\sim}{s}\varepsilon C} | Q_n(\underset{\sim}{s}) | > \varepsilon\} \Rightarrow \{\sup_{\underset{\sim}{s}\varepsilon C_u} | Q_n(\underset{\sim}{s}) | > \varepsilon\} \text{ for}$$

some $u = 1, 2, \ldots, U$. It suffices then to show that each term in the above finite sum converges to zero, as $n\to\infty$. But,

$$P[\sup_{\underset{\sim}{s}\epsilon C_u} | Q_n(\underset{\sim}{s}) | > \epsilon] = P[\sup_{\underset{\sim}{s}\epsilon C_u} | Q_{n,u}(\underset{\sim}{s}) + Q_{n,o}(\underset{\sim}{s}_u) | > \epsilon]$$

$$\leq P[\sup_{\underset{\sim}{s}\epsilon C_u} | Q_{n,u}(\underset{\sim}{s}) | > \tfrac{\epsilon}{2}] + P[| Q_{n,o}(\underset{\sim}{s}_u) | > \tfrac{\epsilon}{2}] .$$

We shall next apply Lemma 3.2.3 to show that each of the probabilities on the right hand side of the above inequality converges to zero. First consider

$$Q_{n,o}(\underset{\sim}{s}_u) = \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\epsilon A} n^{1/2} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\, \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\, \underset{\sim}{\beta})]$$

Applying lemma 3.2.3 with

$$\tilde{k}_n(\cdot) = n^{1/2} [\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \tilde{h}_n(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\beta})] ,$$

shows $Q_{n,o}(s_u) \overset{P}{\to} 0$, provided we can show that

$$\lim_{n\to\infty} E[\{\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \tilde{h}_n(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\beta})\}^2] = 0 .$$

But

$$\{\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \tilde{h}_n(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})\}^2$$

$$= \{h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)$$

$$- h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})\}^2$$

$$= \{h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})$$

$$- [\theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \theta_n(\underset{\sim}{\beta})]\}^2$$

$$\leq 2[h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})]^2$$

$$+ 2[\theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \theta_n(\underset{\sim}{\beta})]^2$$

Taking expectations, we have

$$E\{[\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \tilde{h}_n(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})]^2\}$$

$$\leq 2E\{[h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})]^2\}$$

$$+ 2[\theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - \theta_n(\underset{\sim}{\beta})]^2$$

$$\leq 4E[|h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})|^2]$$

$$\leq 4E[\operatorname*{Sup}_{\underset{\sim}{s}\varepsilon C_u} |h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}) - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})|^2]$$

which goes to zero, as n goes to infinity by (3.2.7) of condition 3.2.5. Here we use the fact that

$$\|\underset{\sim}{\beta} + n^{-1/2}\underset{\sim}{s} - \underset{\sim}{\beta}\| = \|n^{-1/2}\underset{\sim}{s}\| \leq 2n^{-1/2}U\|Cu\| \leq \frac{2\varepsilon U}{8K_1 n^{1/2}} .$$

Next we examine

$$\operatorname*{Sup}_{\underset{\sim}{s}\varepsilon C_u} |Q_{n,u}(\underset{\sim}{s})|$$

$$= \operatorname*{Sup}_{\underset{\sim}{s}\varepsilon C_u} \Big| \frac{n^{1/2}}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} [\tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)] \Big|$$

$$\leq \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} \operatorname*{Sup}_{\underset{\sim}{s}\varepsilon C_u} n^{1/2} \Big| \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) \Big|$$

$$= \frac{1}{\binom{n}{r}} \sum_{\underset{\sim}{\alpha}\varepsilon A} \operatorname{Sup}_{\underset{\sim}{s}\varepsilon C_u} n^{1/2} \Big| \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) \Big|$$

$$- E[\operatorname{Sup}_{\underset{\sim}{s}\varepsilon C_u} n^{1/2} \Big| \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) \Big|$$

$$+ n^{1/2} E\ \{\operatorname{Sup}_{\underset{\sim}{s}\varepsilon C_u} \Big| \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}_n(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) \Big| \}$$

$$\equiv D_{1n} + D_{2n}\ .$$

Now,

$$D_{2n} = n^{1/2} E[\operatorname{Sup}_{\underset{\sim}{s}\varepsilon C_u} \Big| \tilde{h}(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})$$

$$- \tilde{h}(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\ \underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) \Big|]$$

$$= n^{1/2} E[\sup_{\underset{\sim}{s}\varepsilon C_u} | h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)$$

$$- \Theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}) + \Theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u) |]$$

$$\leq n^{1/2} E[\sup_{\underset{\sim}{s}\varepsilon C_u} |h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|]$$

$$+ n^{1/2} \sup_{\underset{\sim}{s}\varepsilon C_u} |\Theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}) - \Theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|$$

$$= n^{1/2} E[\sup_{\underset{\sim}{s}\varepsilon C_u} |h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|]$$

$$+ n^{1/2} \sup_{\underset{\sim}{s}\varepsilon C_u} |E[h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)]|$$

$$\leq 2n^{1/2} E[\sup_{\underset{\sim}{s}\varepsilon C_u} |h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{\alpha_1:n},\ldots,X_{\alpha_r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|]$$

$$\leq 2n^{1/2} K_1 \|C_u\| n^{-1/2} = 2K_1 \|C_u\| \leq \frac{\varepsilon}{4} \text{ by}$$

(3.2.6) of Condition 3.2.5, the definition of C_u, and

$$\|\underset{\sim}{\beta} + n^{-1/2}\underset{\sim}{s} - \underset{\sim}{\beta} - n^{-1/2}\underset{\sim}{s}_u\| = n^{-1/2}\|\underset{\sim}{s} - \underset{\sim}{s}_u\| < n^{-1/2}\|C_u\| .$$

Next consider D_{1n} and apply lemma 3.2.3 with

$$\tilde{K}_n(\cdot)$$

$$= n^{1/2}\{[\sup_{\underset{\sim}{s}\in C_u}|\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|$$

$$- E[\sup_{\underset{\sim}{s}\in C_u}|\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|]\}.$$

Now,

$$\frac{1}{n}E[\{\tilde{K}_n(X_{1:n},\ldots,X_{r:n})\}^2]$$

$$= E\ \{\sup_{\underset{\sim}{s}\in C_u}|\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|$$

$$- E[\sup_{\underset{\sim}{s}\in C_u}|\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|]\}^2$$

$$\leq E\ [\sup_{\underset{\sim}{s}\in C_u}|\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\tilde{h}_n(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)|^2]$$

$$\leq 2E\,[\underset{\underset{\sim}{s}\in C_u}{\text{Sup}}\left|h(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)\right|]$$

$$+\ 2\underset{\underset{\sim}{s}\in C_u}{\text{Sup}}\left|\theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-\theta_n(\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)\right|^2$$

$$\leq 4E\,[\underset{\underset{\sim}{s}\in C_u}{\text{Sup}}\left|h(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s})-h(X_{1:n},\ldots,X_{r:n};\underset{\sim}{\beta}+n^{-1/2}\underset{\sim}{s}_u)\right|^2]$$

$\to 0$, as $n \to \beta$ by (3.2.7) of Condition 3.2.5 and since

$$n^{-1/2}\|C_u\| \to 0\ ,\ \text{as } n \to \infty\ .$$

Thus far, we have shown that under the Conditions 3.2.4 and 3.2.5

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}})-\theta_n(\hat{\underset{\sim}{\beta}})-U_n(\underset{\sim}{\beta})+\theta_n(\underset{\sim}{\beta})] \overset{P}{\to} 0\ ,\ \text{as } n\to\infty\ .$$

The main result of this section is given in the next theorem which yields the limiting distribution of

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}})-\theta_n(\underset{\sim}{\beta})]\ .$$

<u>THEOREM 3.2.10</u>

Suppose that $\theta_n(\underset{\sim}{\gamma})$ is uniformly (in n) differentiable at $\underset{\sim}{\gamma}=\underset{\sim}{\beta}$ and that this differential is zero. Suppose further that the conditions of Theorem 3.2.8 are satisfied. If, in addition,

$$n^{1/2}[U_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta})] \overset{d}{\to} N(0,\tau^2) \text{ , as } n\to\infty \text{ ,} \tag{3.2.11}$$

with $\tau^2 > 0$, then

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] \overset{d}{\to} N(0,\tau^2) \text{ .}$$

<u>PROOF.</u> Note that

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})]$$

$$= n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\hat{\underset{\sim}{\beta}}) - U_n(\underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})]$$

$$+ n^{1/2}[U_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta}) - \theta_n(\hat{\underset{\sim}{\beta}}) + \theta_n(\underset{\sim}{\beta})]$$

$$= n^{1/2}[U_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta})] + n^{1/2}[\theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] + o_p(1) \text{ ,}$$

since by Theorem 3.2.8

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\hat{\underset{\sim}{\beta}}) - U_n(\underset{\sim}{\beta}) + \theta_n(\underset{\sim}{\beta})] \overset{P}{\to} 0 \text{ , as } n\to\infty \text{ .}$$

Then, by Slutsky's theorem,

$$n^{1/2}[U_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] \quad \text{and} \quad n^{1/2}[U_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta})]$$

have the same limiting distribution, provided that we can show that

$$n^{1/2}[\theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] = o_p(1) . \tag{3.2.12}$$

We show below that this follows from the fact that $\theta_n(\underset{\sim}{\gamma})$ is uniformly differentiable at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$, and that this differential is zero at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$. By definition (see, for example, Serfling, 1980, p. 45), $\theta_n(\underset{\sim}{\gamma})$ is <u>uniformly</u> (in n) <u>differentiable</u> at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$ if for every n, $(\partial\theta_n)/(\partial\gamma_i)$, i = 1, 2, . . . , p, all exist and if, in addition, the <u>differential function</u>

$$\sum_{i=1}^{p} \frac{\partial\theta_n}{\partial\gamma_i}\bigg|_{\underset{\sim}{\gamma} = \underset{\sim}{\beta}} \cdot(\gamma_i - \beta_i)$$

satisfies the property that for every $\varepsilon>0$, there exists a neighborhood $N_\varepsilon(\underset{\sim}{\beta})$ of $\underset{\sim}{\beta}$ and an N_ε^* such that for every $\underset{\sim}{\gamma}\in N_\varepsilon(\underset{\sim}{\beta})$ and for $n>N_\varepsilon^*$

$$\left| \theta_n(\underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\beta}) - \sum^{p} (\gamma_i - \beta_i) \frac{\partial\theta_n}{\partial\gamma_i}\bigg|_{\underset{\sim}{\gamma} = \underset{\sim}{\beta}} \right| \leq \varepsilon \, \|\underset{\sim}{\gamma} - \underset{\sim}{\beta}\| .$$

Now since θ_n admits a zero differential at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$, and since it is uniformly (in n) differentiable at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$ we have that for every $\varepsilon>0$ there exists $N_\varepsilon(\underset{\sim}{\beta})$, a neighborhood of $\underset{\sim}{\beta}$, and N_ε^* such that

$$\left| \theta_n(\underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\beta}) \right| \leq \varepsilon \, \|\underset{\sim}{\gamma} - \underset{\sim}{\beta}\| ,$$

whenever $\underset{\sim}{\gamma}\in N_\varepsilon(\underset{\sim}{\beta})$ and $n>N_\varepsilon^*$. It follows that for $\hat{\underset{\sim}{\beta}}$ in $N_\varepsilon(\underset{\sim}{\beta})$, and $n>N_\varepsilon^*$

$$n^{1/2} \left| \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \right| \leq \varepsilon n^{1/2} \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| . \qquad (3.2.13)$$

Also, by Condition 3.2.4,

$$n^{1/2} (\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}) = O_p(1)$$

which implies that

$$n^{1/2} \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| = O_p(1) \qquad (3.2.14)$$

since

$$n^{1/2} \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| = n^{1/2} [\sum_{i=1}^{p} (\hat{\beta}_i - \beta_i)^2]^{1/2}$$

$$\leq \sum_{i=1}^{p} | n^{1/2} (\hat{\beta}_i - \beta_i) | ,$$

and since a finite sum of $O_p(1)$ variables is $O_p(1)$. By (3.2.14), we know that for every $\delta>0$ there exists $M_\delta>0$ such that

$$P \{n^{1/2} \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| > M_\delta\} < \delta , \qquad (3.2.15)$$

for every n. Now, to show (3.2.12) we need to show that for every $\varepsilon^*>0$ and every $\delta^*>0$, there exists an N such that

$$P\{n^{1/2} \mid \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \mid > \varepsilon^*\} < \delta^* ,$$

whenever $n > N$.

Take $\delta = \delta^*/2$ and let $\varepsilon = \varepsilon^*/M_\delta$ where M_δ is defined by (3.2.15). By (3.2.13) we know that there exists a neighborhood of $\underset{\sim}{\beta}$ with radius d_ε, and there exists an N_ε^* such that $n>N_\varepsilon^*$ and $\|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\|<d_\varepsilon$ imply

$$n^{1/2} \mid \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \mid \le \frac{\varepsilon^*}{M_\delta} n^{1/2} \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| .$$

Choose N_1 so that $n>N_1$ implies

$$P\{\|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| > d_\varepsilon\} < \frac{\delta^*}{2} .$$

(Note that the choice of such an n is possible since $n^{1/2}\|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| = 0_p(1)$). Combining the above observations we see that for every $\varepsilon^*>0$, every $\delta^*>0$ and for $n>N \equiv \max(N_\varepsilon^*, N_1)$,

$$P\{n^{1/2} \mid \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \mid > \varepsilon^*\}$$

$$= P\{n^{1/2} \mid \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \mid > \varepsilon^* ,\ \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| > d_\varepsilon\}$$

$$+ P\{n^{1/2} \mid \theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta}) \mid > \varepsilon^* ,\ \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| \le d_\varepsilon\}$$

$$\leq \frac{\delta^*}{2} + P\,\{n^{1/2}\,|\,\theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})\,| > \varepsilon^*\,,\ \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| \leq d_\varepsilon\}$$

$$\leq \frac{\delta^*}{2} + P\,\{\frac{\varepsilon^*}{M_\delta}\, n^{1/2}\, \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| > \varepsilon^*\}$$

$$= \frac{\delta^*}{2} + P\,\{\, n^{1/2}\, \|\hat{\underset{\sim}{\beta}} - \underset{\sim}{\beta}\| > M_\delta\,\}$$

$$\leq \frac{\delta^*}{2} + \frac{\delta^*}{2} = \delta^*\,,$$

which implies that $n^{1/2}\,|\theta_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})| = o_p(1)$.

REMARK 3.2.16

Although the above results deal with a random sample of observations from a univariate distribution, they remain valid when $\underset{\sim}{X}_{1:n}, \ldots, \underset{\sim}{X}_{n:n}$ come from some multivariate population.

REMARK 3.2.17

A difficult step in applying theorem 3.2.10 is verifying (3.2.7) of Condition 3.2.5. However, if one can show that there exists an $M_1>0$ such that

$$|\, h(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\gamma}) - h(X_{1:n},\ldots,X_{r:n};\, \underset{\sim}{\beta})\,| \leq M_1 \qquad (3.2.18)$$

for all $\underset{\sim}{\gamma}$ in some neighborhood of $\underset{\sim}{\beta}$, and every $X_{1:n}, \ldots, X_{r:n}$, then (3.2.6) implies (3.2.7). To see this note that

$$E\Big[\sup_{\underset{\sim}{\gamma}'\varepsilon D(\underset{\sim}{\gamma},d)} |h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\gamma}') - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})|^2\Big]$$

$$= E\Big[\Big\{\sup_{\underset{\sim}{\gamma}'\varepsilon D(\underset{\sim}{\gamma},d)} |h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\gamma}') - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})|\Big\}^2\Big]$$

$$\leq M_1 E\Big[\sup_{\underset{\sim}{\gamma}'\varepsilon D(\underset{\sim}{\gamma},d)} |h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\gamma}') - h(X_{1:n},\ldots,X_{r:n};\ \underset{\sim}{\beta})|\Big]$$

$$\leq M_1 K_1 d$$

which goes to zero as $d \to 0$ by (3.2.6).

3.3 The Asymptotic Normality of T_n Under a Sequence of Alternatives

The statistic T_n involves the estimator $\hat{\underset{\sim}{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ of $\underset{\sim}{\beta} = (\beta_1, \beta_2)'$, so that to apply theorem 3.2.10 we need first to obtain the asymptotic normality of $n^{1/2}[T_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta})]$ under a sequence of alternatives, i.e., we need the asymptotic normality of the statistic involving the parameter value $\underset{\sim}{\beta}$ rather than its estimate $\hat{\underset{\sim}{\beta}}$. To this end, we shall apply theorem 5.3.10, and lemmas 5.3.11 and 5.3.13 of Randles and Wolfe (1979).

Using the "trivariate reduction" method, we may write

$$E_i = W_{1i} + \Delta W_{3i}$$

and

$$E_i' = W_{2i} + \Delta W_{3i}\ ,\ i=1,2,\ldots,n\ ,$$

so that our underlying linear models are given by

$$Y_i = \alpha_1 + \beta_1 X_i + W_{1i} + \Delta W_{3i} \text{ , and}$$

$$Z_i = \alpha_2 + \beta_2 X_i + W_{2i} + \Delta W_{3i} \text{ , } i=1,2,\ldots,n \text{ ,}$$

and the hypothesis of independence of E_i and E_i', $i = 1, 2, \ldots, n$, is <u>equivalent</u> to the hypothesis $\Delta = 0$. To establish the results of this section, we need the following assumptions:

<u>3.3.1</u> $\{W_{1i}\}$, $\{W_{2i}\}$ and $\{W_{3i}\}$, $i = 1, 2, \ldots, n$, are three independent random samples of random variables with absolutely continuous distribution functions $G_1(.)$, $G_2(.)$ and $G_3(.)$, respectively.

<u>3.3.2</u> The variables $T_k = W_{k1} - W_{k2}$ have distribution functions $F_k(.)$ and bounded and continuous density functions $f_k(.)$, $k = 1, 2, 3$.

<u>3.3.3</u> The variable X_1 has a finite first moment.

Let $\underset{\sim}{S}_{1:n}, \ldots, \underset{\sim}{S}_{n:n}$ denote a random sample from some trivariate distribution with distribution function $G_n(.,.,.)$ depending on n, where

$$\underset{\sim}{S}_{i:n} = \begin{bmatrix} X_i \\ W_{1i} + \Delta_n W_{3i} \\ W_{2i} + \Delta_n W_{3i} \end{bmatrix} . \qquad (3.3.4)$$

The symmetric kernel of degree r=2 is then given by

$$h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\gamma}) = \text{Sgn}\ \{[(W_{11}-W_{12})+\Delta_n(W_{31}-W_{32})-(\gamma_1-\beta_1)(X_1-X_2)]$$
$$\cdot[(W_{21}-W_{22})+\Delta_n(W_{31}-W_{32})-(\gamma_2-\beta_2)(X_1-X_2)]\}$$

where Δ_n is a sequence of parameters depending on n with $\Delta_n \to 0$, as $n\to\infty$, $\underset{\sim}{\beta} = (\beta_1,\beta_2)'$ is a fixed parameter and $\underset{\sim}{\gamma} = (\gamma_1,\gamma_2)'$ is a mathematical variable.

The kernel may be rewritten as

$$h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\gamma}) = \text{Sgn}\ \{[T_1+\Delta_n T_3-(\gamma_1-\beta_1)(X_1-X_2)]$$
$$\cdot[T_2+\Delta_n T_3-(\gamma_2-\beta_2)(X_1-X_2)]\}\ ,$$

and the corresponding U-statistic is

$$T_n(\underset{\sim}{\gamma}) = T_n(\underset{\sim}{S}_{1:n},\ldots,\underset{\sim}{S}_{n:n};\ \underset{\sim}{\gamma})$$
$$= \frac{1}{\binom{n}{2}} \sum_{i<j} h(\underset{\sim}{S}_{i:n},\underset{\sim}{S}_{j:n};\ \underset{\sim}{\gamma})\ .$$

To obtain the asymptotic distribution of $T_n(\underset{\sim}{\beta})$, we first need to find its mean and its limiting variance, which we shall do next. Note that

$$\theta_n(\underset{\sim}{\beta}) = E[T_n(\underset{\sim}{\beta})]$$
$$= E[h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\beta})]$$
$$= E[\text{Sgn}\{[T_1+\Delta_n T_3][T_2+\Delta_n T_3]\}]$$

$$= P\{[T_1+\Delta_n T_3][T_2+\Delta_n T_3] > 0\}$$

$$- P\{[T_1+\Delta_n T_3][T_2+\Delta_n T_3] < 0\}$$

$$= P\{T_1+\Delta_n T_3 > 0,\ T_2+\Delta_n T_3 > 0\}$$

$$+ P\{T_1+\Delta_n T_3 < 0,\ T_2+\Delta_n T_3 < 0\}$$

$$- P\{T_1+\Delta_n T_3 > 0,\ T_2+\Delta_n T_3 < 0\}$$

$$- P\{T_1+\Delta_n T_3 < 0,\ T_2+\Delta_n T_3 > 0\}$$

$$= 2P\{T_1+\Delta_n T_3 > 0,\ T_2+\Delta_n T_3 > 0\}$$

$$+ 2P\{T_1+\Delta_n T_3 < 0,\ T_2+\Delta_n T_3 < 0\} - 1$$

$$= E_{T_3}\ \{2[1-F_1(-\Delta_n T_3)][1-F_2(-\Delta_n T_3)]+2F_1(-\Delta_n T_3)F_2(-\Delta_n T_3)-1\} \qquad (3.3.5)$$

where E_{T_3} denotes expectation with respect to the random variable T_3.

The asymptotic variance of $T_n(\underset{\sim}{\beta})$ is $\eta = \lim_{n\to\infty} r^2\zeta_{1:n}$ with

$$\zeta_{1:n} = \text{Var}\ [h_{1:n}(\underset{\sim}{S}_{1:n})]$$

and

$$h_{1:n}(\underset{\sim}{s}) = E\ [h(\underset{\sim}{s},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\beta})\ |\ \underset{\sim}{S}_{1:n} = \underset{\sim}{s}]\ .$$

The limiting variance, η, will be obtained as a result of applying theorem 5.3.10 of Randles and Wolfe (1979), for which we only need the quantity $h_{1:n}(\underset{\sim}{s})$. We see that

$$h_{1:n}(\underset{\sim}{s}) = E\,[h(\underset{\sim}{s},\underset{\sim}{S}_{2:n};\,\underset{\sim}{\beta})\mid \underset{\sim}{S}_{1:n} = \underset{\sim}{s} = (x,e_n,e_n')']$$

$$= E[\mathrm{Sgn}\{[(W_{11}-W_{12})+\Delta_n(W_{31}-W_{32})][(W_{21}-W_{22})+\Delta_n(W_{31}-W_{32})]\mid \underset{\sim}{S}_{1:n}= \underset{\sim}{s}\}]$$

$$= E[\mathrm{Sgn}\{[e_n-W_{12}-\Delta_nW_{32}][e_n'-W_{22}-\Delta_nW_{32}]\}]$$

$$= 2P\{e_n-W_{12}-\Delta_nW_{32} > 0,\ e_n'-W_{22}-\Delta_nW_{32} > 0\}$$

$$+\ 2P\{e_n-W_{12}-\Delta_nW_{32} < 0,\ e_n'-W_{22}-\Delta_nW_{32} < 0\} - 1$$

$$= E_{W_{32}}\ \{2G_1(e_n-\Delta_nW_{32})G_2(e_n'-\Delta_nW_{32})$$

$$+\ 2[1-G_1(e_n-\Delta_nW_{32})][1-G_2(e_n'-\Delta_nW_{32})] -1\}\ , \qquad (3.3.6)$$

where $G_1(.)$ $[G_2(.)]$ is the distribution function of $W_{12}[W_{22}(.)]$, and e_n (e_n') is of the form $w_1+\Delta_nw_3$ $(w_2+\Delta_nw_3)$ with w_1, w_2 and w_3 being given values of W_{11}, W_{21} and W_{31}.

Next, we verify the conditions of theorem 5.3.10 of Randles and Wolfe. Condition (i) is immediate, since

$$E[h^2(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n})] = 1\ ,\ \text{for every } n \geq 2\ . \qquad (3.3.7)$$

Conditions (ii) and (iii) hold, if the conditions of lemmas 5.3.11 and 5.3.13 are satisfied. Lemma 5.3.11 follows from (3.3.7) with $M = 1$. There remains to verify conditions (i) - (iv) of lemma 5.3.13.

<u>Condition (i)</u>: We need to show that there exists a real valued

function $k(\underset{\sim}{s})$ such that

$$\lim_{n\to\infty} h_{1:n}(\underset{\sim}{s}) = k(\underset{\sim}{s}) \quad \text{for every } \underset{\sim}{x} .$$

But from (3.3.5), and for every $\underset{\sim}{s} = (x,e_n,e_n')'$,

$$\begin{aligned}
&\lim_{n\to\infty} h_{1:n}(\underset{\sim}{s}) \\
&= \lim_{n\to\infty} E_{W_{32}} \{2G_1(e_n-\Delta_n W_{32})G_2(e_n'-\Delta_n W_{32}) \\
&\qquad + 2[1-G_1(e_n-\Delta_n W_{32})][1-G_2(e_n'-\Delta_n W_{32})] - 1\} \\
&= 2G_1(w_1)G_2(w_2) + 2[1-G_1(w_1)][1-G_2(w_2)] - 1 \\
&\equiv k(\underset{\sim}{s}) ,
\end{aligned}$$

because by the Lebesgue Dominated Convergence Theorem the limit may be passed inside the expected value, since $G_1(.)$ and $G_2(.)$ are absolutely continuous distribution functions, and

$$\lim_{n\to\infty} e_n = w_1 \quad \text{and} \quad \lim_{n\to\infty} e_n' = w_2 .$$

<u>Condition (ii)</u>: Let $G_n(\underset{\sim}{s})$ denote the distribution function of $\underset{\sim}{S}_{i:n}$. We will show that there exists a distribution function $G(\underset{\sim}{s})$ such that

$$\lim_{n\to\infty} G_n(\underset{\sim}{s}) = G(\underset{\sim}{s}) \quad \text{for every } \underset{\sim}{s} ,$$

but this is immediate since $\Delta_n \to 0$, as $n \to \infty$, and, therefore,

$$\underset{\sim}{S}_{i:n} = \begin{vmatrix} X_i \\ W_{1i} + \Delta_n W_{3i} \\ W_{2i} + \Delta_n W_{3i} \end{vmatrix}$$

converges in probability and in law to $\underset{\sim}{S}_i = \begin{vmatrix} X_i \\ W_{1i} \\ W_{2i} \end{vmatrix}$.

Here $G(\underset{\sim}{s})$ is the distribution of $\underset{\sim}{S}_i$, $i = 1, 2, \ldots, n$.

Condition (iii): We need to show there exists an M^* such that $|h_{1:n}(\underset{\sim}{s})| < M^*$ for every $\underset{\sim}{x}$, and every $n \geq 2$. But from the definition of the kernel $h(.,.)$, for every $\underset{\sim}{s}$ and every $n \geq 2$

$$|h_{1:n}(\underset{\sim}{s})| = \left| E\,[h(\underset{\sim}{s}, \underset{\sim}{S}_{2:n};\, \underset{\sim}{\beta}) \mid \underset{\sim}{S}_{1:n} = \underset{\sim}{s}] \right|$$

$$\leq 1 ,$$

and Condition (iii) holds with any $M^* > 1$.

Condition (iv): To find $E(k^2(\underset{\sim}{S})]$, where $\underset{\sim}{S}$ is a random variable with distribution function $G(\underset{\sim}{s})$, recall that

$$k(\underset{\sim}{s}) = 2G_1(w_1)G_2(w_2) + 2[1-G_1(w_1)][1-G_2(w_2)] - 1 ,$$

so that

$$k^2(\underset{\sim}{s}) = 4G_1^2(w_1)G_2^2(w_2) + 4[1-G_1(w_1)]^2[1-G_2(w_2)]^2 + 1 \; ,$$

$$+ 8G_1(w_1)G_2(w_2)[1-G_1(w_1)][1-G_2(w_2)] - 4G_1(w_1)G_2(w_2)$$

$$- 4[1-G_1(w_1)][1-G_2(w_2)] \; .$$

Also note that

$$E[k^2(\underset{\sim}{S})] = E[k^2(W_{11},W_{21})]$$

$$= \int\int k^2(w_1,w_2)\, d\, G(\underset{\sim}{s})$$

$$= \int\int k^2(w_1,w_2)\, d\, G_1(w_1)\, d\, G_2(w_2) \; ,$$

since W_{11}, W_{21} and X are independent. Further,

$$\int G_1^2(w_1)dG_1(w_1) = \int G_2^2(w_2)dG_2(w_2) = \frac{1}{3} \; ,$$

$$\int [1-G_1(w_1)]^2 dG_1(w_1) = \int [1-G_2(w_2)]^2 dG_2(w_2) = \frac{1}{3} \; ,$$

$$\int G_1(w_1)[1-G_1(w_1)]dG_1(w_1) = \int G_2(w_2)[1-G_2(w_2)]dG_2(w_2) = \frac{1}{6} \; ,$$

and

$$\int G_1(w_1)dG_1(w_1) = \int G_2(w_2)dG_2(w_2) = \frac{1}{2} \; ,$$

and therefore

$$E[k^2(\underset{\sim}{S})] = \frac{4}{9} + \frac{4}{9} + 1 + \frac{8}{36} - 1 - 1 = \frac{1}{9} < \infty .$$

Thus, the conditions of lemma 5.3.13 are satisfied, and the limiting variance of $T_n(\underset{\sim}{\beta})$ is

$$\eta = r^2 \mathrm{Var}[k(\underset{\sim}{S})] = 4E[k^2(\underset{\sim}{S})] = \frac{4}{9} ,$$

since

$$E[k(S)] = \iint [2G_1(w_1)G_2(w_2) + 2[1-G_1(w_1)][1-G_2(w_2)] - 1] dG_1(w_1) dG_2(w_2)$$

$$= 0 .$$

Thus we have verified all the conditions of Theorem 5.3.10 in Randles and Wolfe (1979). We have thus proved the following.

<u>THEOREM 3.3.8</u>

Under conditions 3.3.1 - 3.3.3,

$$n^{1/2}[T_n(\underset{\sim}{\beta}) - \theta_n(\underset{\sim}{\beta})] \overset{d}{\to} N(0,\tfrac{4}{9})$$

where $\theta_n(\underset{\sim}{\beta})$ is given in (3.3.5).

Let $\hat{\underset{\sim}{\beta}} = (\hat{\beta}_1, \hat{\beta}_2)'$ be an estimator of $\underset{\sim}{\beta} = (\beta_1, \beta_2)'$. We shall apply our Theorem 3.2.10 to obtain the asymptotic normality of

$$n^{1/2}[T_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})]$$

under a sequence of alternatives approaching the null. To that effect, we need first to verify the conditions of Theorem 3.2.8. Condition 3.2.4 is discussed under Condition 2.4.5 of the previous chapter. Also by remark 3.2.17, step (3.2.7) of Condition 3.2.5 holds if (3.2.6) and (3.2.18) are satisfied. But, from the definition of the kernel h(.), (3.2.18) is immediate with $M_1 = 2$. There remains to prove step (3.2.6) of Condition 3.2.5, i.e., we need to show that there exists a constant $K_1>0$ such that for every n,

$$E[\sup_{\gamma' \varepsilon D(\gamma,d)} |h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\gamma}') - h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\gamma})|] \le K_1 d\ .$$

The proof of this step is identical to that given in verifying Condition 2.4.6., except that here $K_1(.)$ $[k_1(.)]$ and $K_2(.)$ $[k_2(.)]$ denote the distribution functions (density functions) of

$$\frac{T_1+\Delta_n T_3}{X_1-X_2} \quad \text{and} \quad \frac{T_2+\Delta_n T_3}{X_1-X_2}\ ,$$

respectively.

Thus, the conditions of theorem 3.2.8 hold, and to apply theorem 3.2.10 we only need to show that (i) for every n, $\theta_n(\underset{\sim}{\gamma})$ has a zero differential at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$, and that (ii) $\theta_n(\underset{\sim}{\gamma})$ is uniformly (in n) differentiable at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$. Using the notation developed above, we have,

$$\begin{aligned}\theta_n(\underset{\sim}{\gamma}) &= E_{\underset{\sim}{\beta}}[h(\underset{\sim}{S}_{1:n},\underset{\sim}{S}_{2:n};\ \underset{\sim}{\gamma})] \\ &= P\{[T_1+\Delta_n T_3-(\gamma_1-\beta_1)(X_1-X_2)][T_2+\Delta_n T_3-(\gamma_2-\beta_2)(X_1-X_2)] > 0\} \\ &\quad - P\{[T_1+\Delta_n T_3-(\gamma_1-\beta_1)(X_1-X_2)][T_2+\Delta_n T_3-(\gamma_2-\beta_2)(X_1-X_2)] < 0\}\end{aligned}$$

$$= 2P\{T_1+\Delta_n T_3-(\gamma_1-\beta_1)(X_1-X_2)>0,\ T_2+\Delta_n T_3-(\gamma_2-\beta_2)(X_1-X_2)>0\}$$
$$+\ 2P\{T_1+\Delta_n T_3-(\gamma_1-\beta_1)(X_1-X_2)<0,\ T_2+\Delta_n T_3-(\gamma_2-\beta_2)(X_1-X_2)<0\}$$
$$-\ 1\ .$$

Conditioning on X_1, X_2 and T_3, and using the independence of T_1 and T_2, we can write, with $b_i(\underset{\sim}{\gamma}) = (\gamma_i-\beta_i)(X_1-X_2)$, i = 1, 2,

$$\begin{aligned}\theta_n(\underset{\sim}{\gamma}) &= E_{X_1,X_2,T_3}\ \{2P[T_1>b_1(\underset{\sim}{\gamma})-\Delta_n t_3]P[T_2>b_2(\underset{\sim}{\gamma})-\Delta_n t_3]\\ &\qquad + 2P[T_1<b_1(\underset{\sim}{\gamma})-\Delta_n t_3]P[T_2<b_2(\underset{\sim}{\gamma})-\Delta_n t_3]\\ &\qquad - 1 \mid X_1 = x_1,\ X_2 = x_2,\ T_3 = t_3\ \}\\ &= E_{X_1,X_2,T_3}\ \{2[1-F_1(b_1(\underset{\sim}{\gamma})-\Delta_n t_3)][1-F_2(b_2(\underset{\sim}{\gamma})-\Delta_n t_3)]\\ &\qquad + 2F_1(b_1(\underset{\sim}{\gamma})-\Delta_n t_3)F_2(b_2(\underset{\sim}{\gamma})-\Delta_n t_3)\\ &\qquad - 1 \mid X_1 = x_1,\ X_2 = x_2,\ T_3 = t_3\ \} \qquad (3.3.9)\\ &\equiv E_{X_1,X_2,T_3}\ \{J(X_1,X_2,T_3;\ \gamma)\}\ .\end{aligned}$$

By Conditions 3.3.2 and 3.3.4 we can pass differentiation with respect to $\underset{\sim}{\gamma}$ inside the expectation to obtain

$$\frac{\partial \Theta_n(\underset{\sim}{\gamma})}{\partial \underset{\sim}{\gamma}} = E_{X_1,X_2,T_3}\left[\frac{\partial J(X_1,X_2,T_3;\ \underset{\sim}{\gamma})}{\partial \underset{\sim}{\gamma}}\right] .$$

Differentiating first with respect to γ_1, we have

$$\frac{\partial J(.;\ \underset{\sim}{\gamma})}{\partial \gamma_1} = 2(x_1-x_2)f_1(b_1(\underset{\sim}{\gamma})-\Delta_n t_3)[1-2F_2(b_2(\underset{\sim}{\gamma})-\Delta_n t_3)]$$

which, when evaluated at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$, gives

$$2(x_1-x_2)f_1(-\Delta_n t_3)[1-2F_2(-\Delta_n t_3)] .$$

Similarly,

$$\left.\frac{\partial J(.;\ \underset{\sim}{\gamma})}{\partial \gamma_2}\right|_{\underset{\sim}{\gamma} = \underset{\sim}{\beta}} = 2(x_1-x_2)f_2(-\Delta_n t_3)[1-2F_1(-\Delta_n t_3)] .$$

Each of the above two expressions has a zero expectation with respect to X_1 and X_2, so that, for every n, $\Theta_n(\underset{\sim}{\gamma})$ has a zero differential at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$.

To show that $\Theta_n(\underset{\sim}{\gamma})$ is uniformly (in n) differentiable at $\underset{\sim}{\gamma} = \underset{\sim}{\beta}$, we need to show that for every $\varepsilon>0$ there exists $N_\varepsilon(\underset{\sim}{\beta})$, a neighborhood of $\underset{\sim}{\beta}$, and N^*_ε such that for $n>N^*_\varepsilon$,

$$|\Theta_n(\underset{\sim}{\gamma}) - \Theta_n(\underset{\sim}{\beta})| \leq \varepsilon \,\|\underset{\sim}{\gamma} - \underset{\sim}{\beta}\| .$$

To establish this, we use the following lemma which follows immediately by the Lebesque Dominated Convergence Theorem.

LEMMA 3.3.10

Let V denote some random variable (not necessarily independent of X_1 and X_2). If $E[|X_1-X_2|]<\infty$, then

$$E_{X_1,X_2}\{|X_1-X_2|\ |2F(\Delta V)-1|\} \to 0 \text{ as } \Delta\to 0,$$

where F is an absolutely continuous distribution function with $F(0)=\frac{1}{2}$.

Now, from (3.3.5) and (3.3.9)

$$\Theta_n(\underset{\sim}{\beta}) = E_{T_3}\{2[1-F_1(-\Delta_n T_3)][1-F_2(-\Delta_n T_3)]+2F_1(-\Delta_n T_3)F_2(-\Delta_n T_3)-1\}$$

and

$$\Theta_n(\underset{\sim}{\gamma}) = E_{X_1,X_2,T_3}\{2[1-F_1(b_1-\Delta_n T_3)][1-F_2(b_2-\Delta_n T_3)]$$

$$+ 2F_1(b_1-\Delta_n T_3)F_2(b_2-\Delta_n T_3) - 1\},$$

where $b_i = (\gamma_i-\beta_i)(X_1-X_2)$, $i = 1, 2$.

It follows that

$$\theta_n(\underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\beta})$$

$$= E_{X_1,X_2,T_3}\ \{2[1-F_1(b_1-\Delta_n T_3)][1-F_2(b_2-\Delta_n T_3)]$$

$$+ 2F_1(b_1-\Delta_n T_3)F_2(b_2-\Delta_n T_3)$$

$$- 2[1-F_1(-\Delta_n T_3)][1-F_2(-\Delta_n T_3)]$$

$$- 2F_1(-\Delta_n T_3)F_2(-\Delta_n T_3)\ .$$

Subtracting, then adding the quantities

$$2[1-F_1(b_1-\Delta_n T_3)][1-F_2(-\Delta_n T_3)]$$

and

$$2F_1(b_1-\Delta_n T_3)F_2(-\Delta_n T_3)$$

and combining terms, we obtain

$$\theta_n(\underset{\sim}{\gamma}) - \theta_n(\underset{\sim}{\beta})$$

$$= E_{X_1,X_2,T_3}\ \{2[F_2(b_2-\Delta_n T_3)-F_2(-\Delta_n T_3)][2F_1(b_1-\Delta_n T_3)-1]$$

$$+ 2[F_1(b_1-\Delta_n T_3)-F_1(-\Delta_n T_3)][2F_2(-\Delta_n T_3)-1]$$

$$\equiv E_{X_1,X_2,T_3}\ \{Q_1\} + E_{X_1,X_2,T_3}\ \{Q_2\}\ ,$$

where Q_1 and Q_2 are the two terms in the above expectation. Using Taylor's expansion, we have

$$|E_{X_1,X_2,T_3}\{Q_1\}|$$

$$\leq E_{X_1,X_2,T_3}\{|2b_2f_2(b_2-\Delta_nT_3)||2F_1(b_1-\Delta_nT_3)-1|\}$$

$$\leq E_{X_1,X_2,T_3}\{2B_2|b_2||2F_1(b_1-\Delta_nT_3)-1|\}$$

$$\leq 2B_2\|\underset{\sim}{\gamma}-\underset{\sim}{\beta}\|\ E\ \{|X_1-X_2||2F_1(b_1(\underset{\sim}{\gamma})-\Delta_nT_3)-1\}\ ,$$

where

$$b_i \equiv b_i(\underset{\sim}{\gamma}) = (\gamma_i-\beta_i)(X_1-X_2)\ .$$

But, for $\underset{\sim}{\gamma}$ close enough to $\underset{\sim}{\beta}$ and Δ_n sufficiently small, i.e., n sufficiently large, Lemma 3.3.10 shows that we can bound

$$2B_2E\{|X_1-X_2||2F_1(b_1(\underset{\sim}{\gamma})-\Delta_nT_3)-1|\}$$

by $\varepsilon/2$ so that

$$|E_{X_1,X_2,T_3}\{Q_1\}| \leq \frac{1}{2}\varepsilon\|\underset{\sim}{\gamma}-\underset{\sim}{\beta}\|\ .$$

A similar bound exists for $|E_{X_1,X_2,T_3}\{Q_2\}|$, and the result obtains.

All of the conditions of Theorem 3.2.10 have been verified, and therefore we conclude the following.

THEOREM 3.3.11

Under assumptions 3.3.1 - 3.3.3,

$$n^{1/2}[T_n(\hat{\underset{\sim}{\beta}}) - \theta_n(\underset{\sim}{\beta})] \overset{d}{\rightarrow} N(0,\tfrac{4}{9})\ ,\ \text{as } n\rightarrow\infty.$$

3.4 The Asymptotic Normality of Pearson's Partial Correlation Coefficient

The partial correlation between the variables Y and Z with X held constant is defined to be

$$R_{YZ.X} = \frac{R_{YZ} - R_{YX}R_{ZX}}{[(1-R_{YX}^2)(1-R_{ZX}^2)]^{1/2}} , \qquad (3.4.1)$$

where R_{ab} is the usual product moment correlation between the variables a and b, i.e.,

$$R_{ab} = \frac{S_{ab}}{[S_{aa} \cdot S_{bb}]^{1/2}}$$

with

$$S_{ab} = \sum_{i=1}^{n} (a_i - \bar{a})(b_i - \bar{b}) ,$$

and (3.4.2)

$$S_{aa} = \sum_{i=1}^{n} (a_i - \bar{a})^2 .$$

Suppose that Y and Z are both related to X by the simple linear models

$$Y_i = \alpha_1 + \beta_1 X_i + E_i ,$$

and (3.4.3)

$$Z_i = \alpha_2 + \beta_2 X_i + E_i' , \quad i=1,2,\ldots,n .$$

Letting $\hat{\alpha}_1$ $(\hat{\alpha}_2)$ and $\hat{\beta}_1$ $(\hat{\beta}_2)$ be the OLS estimators of α_1 (α_2) and β_1 (β_2), respectively, we obtain the following residuals

$$U_i = Y_i - \hat{\alpha}_1 - \hat{\beta}_1 X_i$$

$$= (\alpha_1 - \hat{\alpha}_1) - (\hat{\beta}_1 - \beta_1) X_i + E_i \ ,$$

and (3.4.4)

$$V_i = (\alpha_2 - \hat{\alpha}_2) - (\hat{\beta}_2 - \beta_2) X_i + E_i' \ , \quad i=1,2,\ldots,n.$$

It can be shown that, under the linear models (3.4.3), $R_{YZ.X}$ is equal to the partial correlation coefficient between E and E' nolding X constant, i.e., $R_{YZ.X} = R_{EE'.X}$. This statistic may be written as

$$R_{EE'.X} = R_{YZ.X} = R_{UV}$$

$$= \frac{R_{EE'} - R_{EX} \cdot R_{E'X}}{[(1-R^2_{EX})(1-R^2_{E'X})]^{1/2}} \ .$$

Expressing each of the correlation coefficients in the above expression in terms of the appropriate sums of squares and cross products, we have

$$R_{EE'.X} = \frac{S_{XX}S_{EE'} - S_{XE}S_{XE'}}{[(S_{XX}S_{EE} - S_{XE}^2)(S_{XE}S_{E'E'} - S_{XE'}^2)]^{1/2}} \qquad (3.4.5)$$

For efficiency studies, we shall obtain the asymptotic normality of the partial correlation statistic under the "trivariate reduction" model proposed earlier, i.e., when E and E' are related by

$$E_i = W_{1i} + \Delta W_{3i} \quad ,$$

and (3.4.6)

$$E_i' = W_{2i} + \Delta W_{3i} \quad , \; i=1,2,\ldots,n \; ,$$

where $\{W_{1i}\}$, $\{W_{2i}\}$, $\{W_{3i}\}$, i = 1, 2, . . . , n, are three independent random samples having the same distribution as the continuous random variables W_1, W_2 and W_3 with distribution functions $G_1(.)$, $G_2(.)$ and $G_3(.)$, respectively. In addition we need the following assumptions:

3.4.7 The variables X, W_1, W_2, W_3 have zero means, and $\sigma_X^2=1$.

3.4.8 The variables W_1, W_2 and W_3 have finite second moments.

3.4.9 The variable X has a finite fourth moment.

Note that there is no loss of generality in assumption 3.4.7. Since the statistic $R_{EE'.X}$ is a function of "translation invariant" cross products and sums of squares, it is free of the locations of X, W_1, W_2 and W_3, and hence no generality is lost in the zero-mean assumption. Also, $R_{EE'.X}$ is free of σ_X^2, the variance of X, since replacing X_i by X_i/σ_X does not affect the value of $R_{EE'.X}$, so that we may safely take $\sigma_X^2=1$. Denoting the variances of W_1, W_2 and W_3 by σ_1^2, σ_2^2 and σ_3^2, respectively, we see that

$$\mathrm{Corr}(E,E') = \Delta^2\sigma_3^2/[(\sigma_1^2+\Delta^2\sigma_3^2)(\sigma_2^2+\Delta^2\sigma_3^2)]^{1/2} . \qquad (3.4.10)$$

Note again that under the "trivariate reduction" model, the test of independence is equivalent to testing H_0: $\Delta=0$, where in general we may consider Δ to be a function of n. We shall denote the partial correlation coefficient

by $R_n \equiv R_n(\underset{\sim}{X},\underset{\sim}{W}_1,\underset{\sim}{W}_2,\underset{\sim}{W}_3;\ \Delta_n)$, which is the same as the quantity $R_{EE'.X}$, with E and E' being replaced by their corresponding values in terms of Δ and the W's. Using the same notation as in (3.4.2), we calculate the new sums of squares and cross products involved in the statistic R_n to be

$$S_{EE'} = \sum_{i=1}^{n} (E_i-\bar{E})(E_i'-\bar{E}')$$

$$= \sum_{i=1}^{n} [W_{1i}-\bar{W}_1)+\Delta(W_{3i}-\bar{W}_3)][(W_{2i}-\bar{W}_2)+\Delta(W_{3i}-\bar{W}_3)]$$

$$= S_{W_1W_2} + \Delta S_{W_2W_3} + \Delta S_{W_1W_3} + \Delta^2 S_{W_3W_3}$$

and similarly,

$$S_{EE} = S_{W_1W_1} + 2\Delta S_{W_1W_3} + \Delta^2 S_{W_3W_3} ,$$

$$S_{E'E'} = S_{W_2W_2} + 2\Delta S_{W_2W_3} + \Delta^2 S_{W_3W_3} ,$$

$$S_{XE} = S_{XW_1} + \Delta S_{XW_3} , \qquad (3.4.11)$$

and

$$S_{XE'} = S_{XW_2} + \Delta S_{XW_3} \quad .$$

The asymptotic normality of R_n may be obtained in one of two ways. Viewing R_n to be the usual product moment correlation coefficient applied to the residuals, one may think of R_n as a function of three U-statistics, and then use theorems such as that of Randles (1982) or our extended version Theorem 3.2.8 to obtain its asymptotic normality. The other method is to obtain the asymptotic normality of R_n by considering it to be a function of several sample moments. Here, we shall follow the second approach, since it is more straightforward and since it assumes finite moments up to order 4 rather than 6, as would be required by the first approach. For this, we need to apply the following theorem by Kepner (1979):

THEOREM 3.4.12

Let $\underset{\sim}{Q}_{i,n}$ for i = 1, 2, . . . , n be a sequence of n i.i.d. random vectors where

$$\underset{\sim}{Q}_{i,n} = (Q_{1i,n}, \ldots, Q_{pi,n})' ,$$

$$E[\underset{\sim}{Q}_{i,n}] = \underset{\sim}{\mu}_n \quad , \text{ for } i=1,2,\ldots,n ,$$

and

$$\underset{\sim}{\mu}_n \to \underset{\sim}{\mu} \quad \text{as } n\to\infty$$

where

$$\underset{\sim}{\mu}_n = (\mu_{1n}, \ldots, \mu_{pn})' .$$

Let

$$Z_{jn} = \frac{1}{n} \sum_{i=1}^{n} Q_{ji,n} \text{ for } j = 1, 2, \ldots, p$$

and

$$\underset{\sim}{Z}_n = (Z_{1n},\ldots,Z_{pn})' \quad .$$

Let S be a neighborhood of $\underset{\sim}{\mu}$ in R^p and suppose that $g:S \to R$ is a function possessing continuous partial derivatives of order 2 at each point of S. If

$$n^{1/2}[\underset{\sim}{Z}_n - \underset{\sim}{\mu}_n] \overset{d}{\to} N_p(\underset{\sim}{0},\Sigma) ,$$

then

$$n^{1/2}[g(\underset{\sim}{Z}_n)-g(\underset{\sim}{\mu}_n)] \overset{d}{\to} N(0,\underset{\sim}{d}'\Sigma\underset{\sim}{d}) ,$$

where

$$\underset{\sim}{d} = \left[\frac{\partial g(\underset{\sim}{z})}{\partial z_1}\bigg|_{\underset{\sim}{z} = \underset{\sim}{\mu}} ,\ldots, \frac{\partial g(\underset{\sim}{z})}{\partial z_p}\bigg|_{\underset{\sim}{z} = \underset{\sim}{\mu}} \right] \quad .$$

In our case we shall let

$$Q_{1i,n} = X_i \quad , Q_{2i,n} = X_i^2 \quad , Q_{3i,n} = W_{1i} \; ,$$

$$Q_{4i,n} = W_{1i}^2 \quad , Q_{5i,n} = W_{2i} \quad , Q_{6i,n} = W_{2i}^2 \; ,$$

$$Q_{7i,n} = W_{3i} \quad , Q_{8i,n} = W_{3i}^2 \quad , Q_{9i,n} = X_iW_{1i} \; ,$$

$$Q_{10i,n} = X_iW_{2i} \; , Q_{11i,n} = X_iW_{3i} \; , Q_{12i,n} = W_{1i}W_{2i} \; ,$$

$$Q_{13i,n} = W_{1i}W_{3i} \; , Q_{14i,n} = W_{2i}W_{3i} \text{ and } Q_{15i,n} = \Delta_n \; .$$

Suppressing the n subscript on the elements of $\underset{\sim}{Z}_n$, these are given by

$$Z_1 = \bar{X} \;,\; Z_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 \;,\; Z_3 = \bar{W}_1 \;,\; Z_4 = \frac{1}{n}\sum_{i=1}^{n} W_{1i}^2 \;,$$

$$Z_5 = \bar{W}_2 \;,\; Z_6 = \frac{1}{n}\sum_{i=1}^{n} W_{2i}^2 \;,\; Z_7 = \bar{W}_3 \;,\; Z_8 = \frac{1}{n}\sum_{i=1}^{n} W_{3i}^2 \;,$$

$$Z_9 = \frac{1}{n}\sum_{i=1}^{n} X_i W_{1i} \;,\; Z_{10} = \frac{1}{n}\sum_{i=1}^{n} X_i W_{2i} \;,\; Z_{11} = \frac{1}{n}\sum_{i=1}^{n} X_i W_{3i} \;,$$

$$Z_{12} = \frac{1}{n}\sum_{i=1}^{n} W_{1i} W_{2i} \;,\; Z_{13} = \frac{1}{n}\sum_{i=1}^{n} W_{1i} W_{3i} \;,\; Z_{14} = \frac{1}{n}\sum_{i=1}^{n} W_{2i} W_{3i} \;,$$

and $Z_{15} = \Delta_n$, so that

$$\underset{\sim}{\mu}_n = E[\underset{\sim}{Q}_{i,n}] = (0,1,0,\sigma_1^2,0,\sigma_2^2,0,\sigma_3^2,0,0,0,0,0,0,\Delta_n)' \;, \qquad (3.4.13)$$

where $\sigma_i^2 = \mathrm{Var}[W_i]$, i = 1, 2, 3. As functions of $\underset{\sim}{Z}_n$ we can write

$$S_{xx}/n = \sum_{i=1}^{n} X_i^2/n - \bar{X}^2 = Z_2 - Z_1^2$$

and similarly

$$S_{XW_1}/n = Z_9 - Z_1 Z_3 \;,\quad S_{XW_2}/n = Z_{10} - Z_1 Z_5$$

$$S_{XW_3}/n = Z_{11} - Z_1 Z_7 \;,\quad S_{W_1W_1}/n = Z_4 - Z_3^2$$

$$S_{W_1W_2}/n = Z_{12} - Z_3Z_5 \ , \quad S_{W_1W_3}/n = Z_{13} - Z_3Z_7$$

$$S_{W_2W_2}/n = Z_6 - Z_5^2 \ , \ S_{W_2W_3}/n = Z_{14} - Z_5Z_7$$

and, $S_{W_3W_3}/n = Z_8 - Z_7^2$.

It follows that

$$S_{EE'}/n = (Z_{12} - Z_3Z_5) + \Delta(Z_{13} - Z_3Z_7) + \Delta(Z_{14} - Z_5Z_7) + \Delta^2(Z_8 - Z_7^2) \ ,$$

$$S_{EE}/n = (Z_4 - Z_3^2) + 2\Delta(Z_{13} - Z_3Z_7) + \Delta^2(Z_8 - Z_7^2) \ ,$$

$$S_{E'E'}/n = (Z_6 - Z_5^2) + 2\Delta(Z_{14} - Z_5Z_7) + \Delta^2(Z_8 - Z_7^2) \ ,$$

$$S_{XE}/n = (Z_9 - Z_1Z_3) + \Delta(Z_{11} - Z_1Z_7) \ ,$$

and

$$S_{XE'}/n = (Z_{10} - Z_1Z_5) + \Delta(Z_{11} - Z_1Z_7) \ .$$

Substituting in (3.4.11), $R_n \equiv g(\underset{\sim}{Z}_n)$ can be written as

$$g(\underset{\sim}{Z}_n) = \frac{N_1(\underset{\sim}{Z}_n) - N_2(\underset{\sim}{Z}_n)}{D_1^{1/2}(\underset{\sim}{Z}_n)\, D_2^{1/2}(\underset{\sim}{Z}_n)} \qquad (3.4.14)$$

where,

$$N_1(\underset{\sim}{Z}_n) = (Z_2 - Z_1^2)[Z_{12} - Z_3Z_5 + \Delta(Z_{13} - Z_3Z_7) + \Delta(Z_{14} - Z_5Z_7) + \Delta^2(Z_8 - Z_7^2)],$$

$$N_2(\underset{\sim}{Z}_n) = [Z_9 - Z_1Z_3 + \Delta(Z_{11} - Z_1Z_7)][Z_{10} - Z_1Z_5 + \Delta(Z_{11} - Z_1Z_7)]$$

$$D_1(\underset{\sim}{Z}_n) = (Z_2 - Z_1^2)[Z_4 - Z_3^2 + 2\Delta(Z_{13} - Z_3Z_7) + \Delta^2(Z_8 - Z_7^2)]$$

$$- [Z_9 - Z_1Z_3 + \Delta(Z_{11} - Z_1Z_7)]^2 \quad \text{, and}$$

$$D_2(\underset{\sim}{Z}_n) = (Z_2 - Z_1^2)[Z_6 - Z_5^2 + 2\Delta(Z_{14} - Z_5Z_7) + \Delta^2(Z_8 - Z_7^2)]$$

$$- [Z_{10} - Z_1Z_5 + \Delta(Z_{11} - Z_1Z_7)]^2 \quad .$$

With $\underset{\sim}{\mu}_n$ as given in (3.4.13), $N_1(\underset{\sim}{\mu}_n) = \Delta_n^2\sigma_3^2$, $N_2(\underset{\sim}{\mu}_n) = 0$, $D_1(\underset{\sim}{\mu}_n) = \sigma_1^2 + \Delta_n^2\sigma_3^2$ and $D_2(\underset{\sim}{\mu}_n) = \sigma_2^2 + \Delta_n^2\sigma_3^2$, so that

$$g(\underset{\sim}{\mu}_n) = \Delta_n^2\sigma_3^2/[(\sigma_1^2 + \Delta_n^2\sigma_3^2)(\sigma_2^2 + \Delta_n^2\sigma_3^2)]^{1/2} \qquad (3.4.15)$$

which is nothing but Corr(E,E') given in (3.4.10).

Next, define

$$\underset{\sim}{Z}_n^* = (Z_1, Z_2, \ldots, Z_{14})'$$

and

$$\underset{\sim}{\mu}^* = (\mu_1, \mu_2, \ldots, \mu_{14})'$$

where $\mu_1, \mu_2, \ldots, \mu_{14}$ are the first 14 elements of $\underset{\sim}{\mu}_n$ given in (3.4.13), which are free of n. It follows that (see, for example, Serfling, 1980, pp. 125-6)

$$n^{1/2}\left[\underset{\sim}{Z}^{*}_{n} - \underset{\sim}{\mu}^{*}\right] \overset{d}{\rightarrow} N_{14}(0,\Sigma^{*})$$

where Σ^* is the variance-covariance matrix of the vector

$$(X, X^2, W_1, W_1^2, W_2, W_2^2, W_3, W_3^2, XW_1, XW_2, XW_3, W_1W_2, W_1W_3, W_2W_3) .$$

The matrix Σ^* may be written in the partitioned form

$$\Sigma^* = \left[\begin{array}{c|c} M_{8x8} & 0 \\ \hline 0 & S_{6x6} \end{array}\right]$$

where

$$M = \begin{bmatrix} 1 & E[X^3] & 0 & 0 & 0 & 0 & 0 & 0 \\ E[X^3] & E[X^4]-1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_1^2 & E[W_1^3] & 0 & 0 & 0 & 0 \\ 0 & 0 & E[W_1^3] & E[W_1^4]-\sigma_1^4 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_2^2 & E[W_2^3] & 0 & 0 \\ 0 & 0 & 0 & 0 & E[W_2^3] & E[W_2^4]-\sigma_2^4 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & \sigma_3^2 & E[W_3^3] \\ 0 & 0 & 0 & 0 & 0 & 0 & E[W_3^3] & E[W_3^4]-\sigma_3^4 \end{bmatrix} ,$$

and

$$S = \begin{bmatrix} \sigma_1^2 & & & & & 0 \\ & \sigma_2^2 & & & & \\ & & \sigma_3^2 & & & \\ & & & \sigma_1^2\sigma_2^2 & & \\ & 0 & & & \sigma_1^2\sigma_3^2 & \\ & & & & & \sigma_2^2\sigma_3^2 \end{bmatrix} .$$

Also, note that $n^{1/2}[Z_{15n}-\mu_{15n}] = n^{1/2}[\Delta_n-\Delta_n]$ converges in distribution to a normal random variable degenerate at zero, which implies that

$$n^{1/2}[\underset{\sim}{Z}_n - \underset{\sim}{\mu}_n]$$

$$= n^{1/2}\left[\begin{pmatrix} \underset{\sim}{Z}_n^* \\ \Delta_n \end{pmatrix} - \begin{pmatrix} \underset{\sim}{\mu}_n^* \\ \Delta_n \end{pmatrix} \right] \xrightarrow{d} N_{15}(\underset{\sim}{0}, \Sigma)$$

where

$$\Sigma = \left[\begin{array}{c|c} \Sigma^* & \\ \hline & 0 \end{array} \right] .$$

To obtain the asymptotic variance of $R_n \equiv g(\underset{\sim}{Z}_n)$ we need the vector

$$\underset{\sim}{d} = \left\{ \frac{\partial g(\underset{\sim}{z})}{\partial z_i} \Bigg|_{\underset{\sim}{z} = \underset{\sim}{\mu}} \right\}$$

for i = 1, 2, . . . , 15.

Since the 15^{th} diagonal element of $\sum$ is zero, we only need to calculate the elements d_1, d_2, . . . , d_{14} of $\underset{\sim}{d}$. Our calculations indicated that, except for $d_{12} = 1/(\sigma_1\sigma_2)$, the remaining elements of $\underset{\sim}{d}$ are all zero, so that the asymptotic variance of $g(\underset{\sim}{Z}_n)$ is

$$\sigma^2 = \underset{\sim}{d}' \sum \underset{\sim}{d} = 1 .$$

Now, g is a ratio of two polynomial functions whose denominator admits non-zero second order differentials in a neighborhood S of $\underset{\sim}{\mu}$. Therefore, g possesses continuous second order partial derivatives in a neighborhood of $\underset{\sim}{\mu}$ allowing us to apply Theorem 3.4.12 to obtain:

THEOREM 3.4.16

Under conditions 3.4.7 - 3.4.9,

$$n^{1/2}[g(\underset{\sim}{Z}_n) - g(\underset{\sim}{\mu}_n)] \overset{d}{\to} N(0,1),$$

where

$$g(\underset{\sim}{Z}_n) \equiv R_n$$

and

$$g(\underset{\sim}{\mu}_n) = \Delta_n^2\sigma_3^2/[(\sigma_1^2+\Delta_n^2\sigma_3^2)(\sigma_2^2+\Delta_n^2\sigma_3^2)]^{1/2} .$$

3.5 The Pitman Asymptotic Efficiency of T_n Relative to R_n

In this section we shall apply Noether's generalization of a theorem by Pitman to obtain the asymptotic efficiency of T_n relative to R_n, which we shall denote by $ARE(T_n,R_n)$. We first state the theorem by Noether (1955), and then verify its conditions for the two statistics T_n and R_n.

Theorem 3.5.1 (Noether)

Consider testing $H_0:\theta=\theta_0$ versus $H_A:\theta>\theta_0$, let $\{\theta_n\}$ be a sequence of alternative parameters with $\lim_{n\to\infty} \theta_n = \theta_0$.

Suppose the test is based on the statistic $T_n = T(x_1, x_2, \ldots, x_n)$, and let $\Psi_n(\theta)$ and $\sigma_n^2(\theta)$ be functions of θ (in many cases these are respectively the mean and variance of T_n). Assume that

A. $\Psi_n'(\theta_0) = \ldots = \Psi_n^{(m-1)}(\theta_0) = 0 \ , \ \Psi_n^{(m)}(\theta_0) > 0$

B. $\lim_{n\to\infty} n^{-m\delta}\Psi_n^{(m)}(\theta_0)/\sigma_n(\theta_0) = c > 0$, for some $\delta > 0$.

The indicated derivatives are assumed to exist. We shall consider the power of the test based on T_n with respect to the alternative $H':\theta_n=\theta_0+k/n^{\delta}$ where k is an arbitrary positive constant. In addition to A and B assume

C. $\lim_{n\to\infty} \Psi_n^{(m)}(\theta_n)/\Psi_n^{(m)}(\theta_0) = 1$,

and

$\lim_{n\to\infty} \sigma_n(\theta_n)/\sigma_n(\theta_0) = 1$,

and

D. The distribution of $[T_n - \Psi_n(\theta_n)]/\sigma_n(\theta_n)$ tends to the standard normal distribution, both under the alternative hypothesis H' and under the null hypothesis $H_0:\theta_n=\theta_0$. If T_{1n} and T_{2n} are two statistics for testing H_0 against H', and if $m_1=m_2=m$, <u>then</u> the ARE of the two tests satisfying A, B, C and D is given by

$$\lim_{n\to\infty} \frac{R_{2n}^{1/m\delta}(\theta_0)}{R_{1n}^{1/m\delta}(\theta_0)} = \text{ARE}\,(T_{2n},T_{1n}) ,$$

where $R_{in}(\theta) = \Psi_{in}^{(m)}(\theta)/\sigma_{in}(\theta)$, i = 1, 2.

Pitman has called the quantity $R_{in}^{1/m\delta}(\theta_0)$ the <u>efficacy</u> of the i^{th} test in testing the hypothesis $H_0:\theta=\theta_0$.

Our hypothesis is given by $H_0:\Delta=0$ versus $H_A:\Delta>0$, where Δ is such that

$$E_i = W_{1i} + \Delta W_{3i}$$

and

$$E_i' = W_{2i} + \Delta W_{3i} .$$

In addition to assumptions 3.3.1 - 3.3.4 of section 3.3 and assumptions 3.4.7 - 3.4.9 of section 3.4 we need the following assumption:

<u>3.5.2</u> The density functions $f_k(.)$ of $T_k = W_{k1} - W_{k2}$, k = 1, 2, 3, have continuous and bounded derivatives.

Next, we shall verify the conditions of Theorem 3.5.1 for each of the statistics T_n and R_n, using the same notation adopted by Noether. Here, we shall let $\{\Delta_n\}$ denote a sequence of alternative parameters converging to the null, i.e., $\lim_{n\to\infty} \Delta_n = 0$.

Application of 3.5.1 to the statistic T_n:

With $\theta \equiv \Delta$, $\theta_0 \equiv 0$ and $\underset{\sim}{\beta}$ denoting the vector of slope parameters, we have from (3.3.5)

$$\begin{aligned}\Psi_n(\Delta) &= E[\Gamma_n(\underset{\sim}{\beta},\Delta)] \\ &= 2P\{T_1+\Delta T_3 > 0,\ T_2+\Delta T_3 > 0\} + 2P\{T_1+\Delta T_3 < 0,\ T_2+\Delta T_3 < 0\} - 1 \\ &= E_{T_3}\ \{2[1-F_1(-\Delta T_3)][1-F_2(-\Delta T_3)] \\ &\qquad + 2F_1(-\Delta T_3)F_2(-\Delta T_3) - 1\}\ , \qquad (3.5.3)\end{aligned}$$

where $T_k = W_{k1} - W_{k2}$ has distribution function $F_k(.)$ and density $f_k(.)$, $k = 1, 2, 3$, and where E_{T_3} denotes expectation with respect to the variable T_3. Therefore, we can write

$$\Psi_n(\Delta) = \int\{2[1-F_1(-\Delta t)[1-F_2(-\Delta t)]+2F_1(-\Delta t)F_2(-\Delta t)-1\}dF_3(t) \qquad (3.5.4)$$

with $\Psi_n(0) = 0$ since $F_1(0) = F_2(0) = \frac{1}{2}$.

The integrand of the above expression involves continuous bounded functions, so that by theorems such as Theorem A.2.4 of Randles and Wolfe (1979), the derivatives with respect to Δ may be taken inside the integral, to obtain

$$\Psi_n'(\Delta) = \int\{2tf_1(-\Delta t)[1-2F_2(-\Delta t)]+2tf_1(-\Delta t)[1-2F_2(-\Delta t)]\}dF_3(t)$$

and, therefore, $\Psi_n'(0) = 0$ since $F_1(0) = F_2(0) = \frac{1}{2}$.

Using assumption 3.5.2, we differentiate a second time,

$$\Psi_n''(\Delta) = \int\{8t^2f_1(-\Delta t)f_2(-\Delta t) - 2t^2f_1'(-\Delta t)[1-2F_2(-\Delta t)]$$

$$- 2t^2f_2'(-\Delta t)[1-2F_1(-\Delta t)]\}dF_3(t) \ ,$$

so that

$$\Psi_n''(0) = \int 8t^2f_1(0)f_2(0)dF_3(t)$$

$$= 8f_1(0)f_2(0)E[T_3^2]$$

$$= 16\sigma_3^2f_1(0)f_2(0) > 0 \ , \qquad (3.5.5)$$

since with $E[W_3] = 0$, $E[T_3^2] = Var[T_3] = Var[W_{31}-W_{32}] = 2\sigma_3^2$. This satisfies condition A, with m=2. For the remaining conditions we shall take $\sigma_n^2(\Delta_n) \equiv \sigma_n^2(0) \equiv 4/9n$ and $\delta = 1/4$. Condition B follows since

$$\lim_{n\to\infty} n^{-m\delta}\Psi_n^{(m)}(0)/\sigma_n(0)$$

$$= \lim_{n\to\infty} n^{-1/2} \cdot 16\sigma_3^2f_1(0)f_2(0)/(4/9n)^{1/2}$$

$$= 24\sigma_3^2f_1(0)f_2(0) \equiv c > 0 \ .$$

Condition C is immediate from the definition of $\sigma_n(\Delta_n)$. We also see that under assumptions 5.3.2

$$\lim_{n\to\infty} \Psi_n''(\Delta_n) = \Psi''(0) .$$

In sections 2.4 and 3.3, respectively, we have shown that

$$\frac{[T_n - \Psi_n(\Delta_n)]}{\sigma_n(\Delta_n)} \equiv \frac{3}{2} n^{1/2} [T_n - \theta_n(\underset{\sim}{\beta})] \xrightarrow{d} N(0,1) ,$$

as $n\to\infty$, both under the null hypothesis and under a sequence of alternatives, thereby proving condition D. The efficacy of the test based on T_n is then given by

$$R_n^{1/m\delta}(0) = [\Psi_n''(0)/\sigma_n(0)]^2$$

$$= n.576\sigma_3^4 f_1^2(0) f_2^2(0)$$

where $f_1(.)$ and $f_2(.)$ are the probability density functions of $T_1 = W_{11} - W_{12}$ and $T_2 = W_{21} - W_{22}$, respectively, and $\sigma_3^2 = Var[W_3]$.

Application of 3.5.1 to the statistic R_n:

To verify the conditions of Noether's theorem we shall let

$$\Psi_n(\Delta) \equiv \Delta^2\sigma_3^2[(\sigma_1^2+\Delta^2\sigma_3^2)(\sigma_2^2+\Delta^2\sigma_3^2)]^{-1/2}$$

and

$$\sigma_n^2(\Delta) \equiv \sigma_n^2(0) = 1/n ,$$

where $\sigma_i^2 = Var[W_i]$, i = 1, 2, 3. Note that $\Psi_n(0) = 0$, and

$$\Psi_n'(\Delta) = 2\Delta\sigma_3^2(\sigma_1^2+\Delta\sigma_3^2)^{-\frac{1}{2}}(\sigma_2^2+\Delta\sigma_3^2)^{-\frac{1}{2}}$$

$$-\Delta^3\sigma_3^4(\sigma_1^2+\Delta\sigma_3^2)^{-\frac{3}{2}}(\sigma_2^2+\Delta\sigma_3^2)^{-\frac{1}{2}}$$

$$-\Delta^3\sigma_3^4(\sigma_1^2+\Delta\sigma_3^2)^{-\frac{1}{2}}(\sigma_2^2+\Delta\sigma_3^2)^{-\frac{3}{2}},$$

so that $\Psi_n'(0) = 0$. Differentiating a second time and evaluating at $\Delta=0$ we have

$$\Psi_n''(0) = 2\sigma_3^2/\sigma_1\sigma_2 > 0$$

and hence condition A is satisfied with m=2. Condition B is satisfied with m=2 and $\delta = 1/4$, since

$$\lim_{n\to\infty} n^{-m\delta}\Psi_n^{(m)}(0)/\sigma_n(0) = \lim_{n\to\infty} n^{-1/2}.2\sigma_3^2/\sigma_1\sigma_2 n^{-1/2}$$

$$= 2\sigma_3^2/\sigma_1\sigma_2 \equiv c > 0 .$$

Condition C is immediate. Also, in the previous section we have shown that $[R_n-\Psi_n(\Delta_n)]/\sigma_n(\Delta_n)$ converges in distribution to the standard normal distribution, thus obtaining condition D. The efficacy of the test based on R_n is then given by

$$R_n^{1/m\delta}(0) \equiv [\Psi_n''(0)/\sigma_n(0)]^2 = 4n\sigma_3^4/\sigma_1^2\sigma_2^2 .$$

THEOREM 3.5.6

Under assumptions 3.3.1 - 3.3.3, 3.4.7 - 3.4.9 and assumption 3.5.2, the asymptotic efficiency of T_n relative to R_n is

$$ARE(T_n,R_n) = 144\sigma_1^2\sigma_2^2 f_1^2(0) f_2^2(0) \quad . \qquad (3.5.7)$$

Here, $f_1(.)$ $[f_2(.)]$ is the density function of the difference between two i.i.d. random variables $T_1 = W_{11}-W_{12}$ $(T_2 = W_{21}-W_{22})$. Since in the "trivariate reduction" model we implicitly assume knowledge of the distributions of W_1 and W_2, we need to find $f_1(0)$ and $f_2(0)$ in terms of $g_1(.)$ and $g_2(.)$, the respective densities of W_1 and W_2. It can be shown that

$$f_i(0) = \int g_i^2(x)\, dx \ , \ i = 1,2 \ .$$

Using the above relation, we have calculated $ARE(T_n,R_n)$ in the case where W_1, W_2 and W_3 have the same distribution. The results of these calculations for some well known distributions are given in Table 3.1.

Table 3.1

Asymptotic Relative Efficiencies

Distribution	$ARE(T_n,R_n)$
Normal	$9/\pi^2 = 0.912$
Uniform	1
Logistic	1.2
Laplace	2.25

CHAPTER FOUR
THE CORRELATION PROBLEM

4.1 Introduction

Let (Y_1,Z_1), . . . , (Y_n,Z_n) denote a random sample of n observable pairs from some continuous bivariate population with distribution function F. As mentioned in chapter 1, the problem of interest in this chapter is to test the null hypothesis that there is no correlation between the variables Y and Z, versus the alternate hypothesis that a correlation exists between these variables. If we let

$$\tau = P\{(Y_1-Y_2)(Z_1-Z_2)>0\} - P\{(Y_1-Y_2)(Z_1-Z_2)<0\} , \qquad (4.1.1)$$

be the correlation coefficient of interest, the above hypotheses translate to

$$H_0: \ \tau = 0 \quad \text{versus} \quad H_A: \tau \neq 0 , \qquad (4.1.2)$$

or the one-sided alternatives of positive correlation ($\tau>0$) or negative correlation ($\tau<0$). In chapter 1, we discussed the motivation behind using a coefficient such as τ, and hypotheses such as those given in (4.1.2). In particular, we indicated that, at least to us, τ is a most natural measure for a "useful" relationship between the variables, in the sense that its values indicate whether larger values of Y are associated with larger (or smaller) values of Z, and that, therefore, the hypotheses given in (4.1.2) are most appropriate. For these hypotheses, we shall use tests based on Pearson's R and Kendall's tau statistics, although classical tests based on these two statistics assume the null hypothesis of the independence of Y and Z. In section 4.2, we give a brief description of these tests for independence, discuss their properties and their limitations for testing (4.1.2). Although the tests based on Kendall's tau and Pearson's R have different consistency classes ($\tau \neq 0$ for the first, and $\rho \neq 0$ for the second), under the elliptically symmetric models studied in this chapter, these consistency classes are identical, since under such models $\tau \neq 0$ is equivalent to $\rho \neq 0$. We can thus base tests on either R or Kendall's tau without being unfair to either test. In section 4.3, we propose some modifications of these tests in the hope of developing a procedure for testing the null hypothesis that $\tau=0$. Section 4.4 contains the results of a Monte Carlo study investigating the performances of these tests, and our conclusions and recommendations are given in section 4.5.

4.2 Some Tests for Independence

Pearson's product moment correlation coefficient is given by

$$R = \frac{\sum_{i=1}^{n} (Y_i-\bar{Y})(Z_i-\bar{Z})}{\left\{\sum_{i=1}^{n} (Y_i-\bar{Y})^2 \sum_{i=1}^{n} (Z_i-\bar{Z})^2\right\}^{1/2}} \tag{4.2.1}$$

where

$$\bar{Y} = \sum_{i=1}^{n} Y_i/n \qquad \text{and} \qquad \bar{Z} = \sum_{i=1}^{n} Z_i/n \ .$$

The mean of R is

$$E[R] = \rho + O(n^{-1}) \ ,$$

and the variance is given by

$$\text{Var}[R] = \frac{(1-\rho^2)^2}{n} + O(n^{-3/2}) \ , \qquad (4.2.2)$$

where $\rho = \text{Corr}(Y,Z)$. (See, for example, Cramér, 1966, p. 359.) Under the assumption that ρ is 0 and $Y|Z$ (or $Z|Y$) is normal, then

$$T = R[(n-2)/(1-R^2)]^{1/2}$$

has the Student's t-distribution with (n-2) degrees of freedom (see, for example, Anderson, 1958, p. 64). From expression (4.2.2), we note that the asymptotic variance of R depends on the parameter ρ. This motivates the use of a variance-stabilizing transformation. Such a transformation yields what is known as Fisher's Z,

$$Z = \frac{1}{2} \ln \left[(1+R)/(1-R)\right] \ , \qquad (4.2.3)$$

which under the assumption of normality has an limiting mean of

$\frac{1}{2}\ln[(1+\rho)/(1-\rho)]$, and an limiting variance of $1/(n-3)$, so that under the hypothesis of independence ($\rho=0$), $(n-3)^{1/2}Z$ has an asymptotic standard normal distribution. (See, for example, Anderson, 1958, p. 78).

Kendall's tau is a U-statistic estimator of τ given in (4.1.1). It is

$$\hat{\tau} = \frac{1}{\binom{n}{2}} \sum_{i<j} \text{Sgn}\{(Y_i - Y_j)(Z_i - Z_j)\} , \qquad (4.2.4)$$

where

$$\text{Sgn}(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t = 0 \; . \\ -1 & \text{if } t < 0 \end{cases}$$

This U-statistic has a symmetric kernel of degree 2 given by

$$h(\underset{\sim}{X}_1, \underset{\sim}{X}_2) = \text{Sgn}\{(Y_1 - Y_2)(Z_1 - Z_2)\} ,$$

with $\underset{\sim}{X} = (Y,Z)'$. Note that

$$\begin{aligned} E[\hat{\tau}] &= E[h(\underset{\sim}{X}_1, \underset{\sim}{X}_2)] \\ &= P\{(Y_1 - Y_2)(Z_1 - Z_2) > 0\} - P\{(Y_1 - Y_2)(Z_1 - Z_2) < 0\} \\ &= \tau \; . \end{aligned}$$

Using results on the variance of a U-statistic, we have

$$\mathrm{Var}[\hat{\tau}] = \frac{1}{\binom{n}{2}} [2(n-2)\zeta_1 + \zeta_2] , \tag{4.2.5}$$

where

$$\zeta_1 = E[h(\underset{\sim}{X}_1,\underset{\sim}{X}_2) \ . \ h(\underset{\sim}{X}_1,\underset{\sim}{X}_3)] - \tau^2$$

and (4.2.6)

$$\zeta_2 = E[h(\underset{\sim}{X}_1,\underset{\sim}{X}_2) \ . \ h(\underset{\sim}{X}_1,\underset{\sim}{X}_2)] - \tau^2$$

$$= 1 - \tau^2 .$$

Letting $h_1(\underset{\sim}{x})$ denote $E[h(\underset{\sim}{X}_1,\underset{\sim}{X}_2)|\underset{\sim}{X}_1 = \underset{\sim}{x}]$, and noting that $E[h_1(\underset{\sim}{X}_1)] = \tau$, we can write

$$\zeta_1 = E\{E[h(\underset{\sim}{X}_1,\underset{\sim}{X}_2)h(\underset{\sim}{X}_1,\underset{\sim}{X}_3)]|\underset{\sim}{X}_1 = \underset{\sim}{x}\} - \tau^2$$

$$= E[h_1^2(\underset{\sim}{X}_1)] - \tau^2$$

$$= \mathrm{Var}[h_1(\underset{\sim}{X}_1)] . \tag{4.2.7}$$

Under the hypothesis of the independence of Y and Z, $\tau=0$ and $\zeta_1 = 1/9$, so that the variance of $\hat{\tau}$ simplifies to

$$\mathrm{Var}_0[\hat{\tau}] = \frac{2(2n+5)}{9n(n-1)} . \tag{4.2.8}$$

In general, however, Var$[\hat{\tau}]$ depends on the underlying bivariate distribution of (Y,Z).

To compare the powers of the tests based on the statistics R and $\hat{\tau}$, one needs to define a suitable class of alternatives, i.e., a class of alternatives which is reasonably wide and reasonably easy to handle mathematically. One such class of alternatives was formulated by H.S. Konijn (1956). Similar classes were also proposed by S. Bhuchongkul (1964) and D.V. Gokhale (1978). To obtain the class of alternatives, Konijn defines

$$Y = \lambda_1 W_1 + \lambda_2 W_2$$

and

$$Z = \lambda_3 W_1 + \lambda_4 W_2 \quad ,$$

where W_1 and W_2 are two independent random variables, and the hypothesis to be tested is

$$H_0: \ \lambda_2 = \lambda_3 = 0 \ .$$

Konijn reports the asymptotic efficiency of $\hat{\tau}$ relative to R for several distributions, in the case when W_1 and W_2 are identically distributed. The values of these AREs are $9/\pi^2 = 0.92$, 1, 0.86, and 1.266 for the normal, uniform, parabolic ($f(t)=kt^2$, for $a \leq t \leq b$), and the Laplace distributions, respectively. To compare the empirical powers of tests based on the statistics R and $\hat{\tau}$ through a Monte Carlo simulation, we adopted a class of alternatives similar to the one proposed by Konijn, but involving only one parameter, Δ. This class of alternatives was suggested by Hájek and Sidák (1967) and is given by

$$Y = W_1 + \Delta W_3$$

and (4.2.9)

$$Z = W_2 + \Delta W_3 \quad ,$$

with W_1, W_2 and W_3 being mutually independent, so that the hypothesis of independence is equivalent to

$$H_0: \quad \Delta = 0 \; .$$

Based on the AREs reported by Konijn, we expected Kendall's tau to perform better for heavy-tailed distributions. To our surprise, however, we found in our Monte Carlo studies that Pearson's R exhibited a high degree of robustness in terms of its stable empirical α-level and empirical power even for such heavy-tailed distributions such as the Cauchy distribution. To test the broader null hypothesis τ=0, we calculated empirical levels and powers for pairs of observations from some bivariate elliptically symmetric distributions (see Johnson and Ramberg, 1977). Here, the empirical α-levels for tests based on both statistics, R and $\hat{\tau}$, were largely inflated, although the α-levels for tests based on Pearson's R were much higher (details of this and other studies are given in sections 4.4 and 4.5). We suspected that these inflated levels were due to the fact that under H_0: $\tau=0$, the variances of R and $\hat{\tau}$ are different from those under the hypothesis of independence. This and other observations motivated us to propose some modifications to the classical tests based on R and $\hat{\tau}$. A discussion of this is given in the next section.

4.3 Tests for Correlation

If the hypothesis of the independence of Y and Z is relaxed, many of the properties of $\hat{\tau}$ and R discussed in the previous section no longer hold. For example, under the hypothesis that $\tau=0$, $E[\hat{\tau}]=0$, but the variance of $\hat{\tau}$ depends on the underlying distribution F, and hence $\hat{\tau}$ is neither distribution-free nor asymptotically distribution-free (see the expression for $Var[\hat{\tau}]$ given in (1.2.5)). From U-statistic theory, we know that

$$(\hat{\tau} - \tau)/\{Var[\hat{\tau}]\}^{1/2} \xrightarrow{d} N(0,1)$$

and

$$\frac{n^{1/2}(\hat{\tau} - \tau)}{(4\zeta_1)^{1/2}} \xrightarrow{d} N(0,1), \text{ as } n\to\infty .$$

To test the hypothesis $\tau=0$, Fligner and Rust (1983) considered several estimators for $4\zeta_1$. They recommend the use of the jackknife estimator $\hat{\sigma}_J^2$ given by

$$\hat{\sigma}_J^2 = (n-1) \sum_{i=1}^{n} (\hat{\tau}^{(i)} - \hat{\tau})^2 \quad , \tag{4.3.1}$$

where $\hat{\tau}^{(i)}$ is Kendall's tau computed on the subsample of size (n-1) formed by leaving out the (Y_i, Z_i) pair. If one defines C_i as

$$C_i = \sum_{\substack{j=1 \\ j \neq i}}^{n} \mathrm{Sgn}\{(Y_i - Y_j)(Z_i - Z_j)\} \;, \; i=1,2,\ldots,n \;, \tag{4.3.2}$$

then, one can show that

$$\hat{\tau} = \sum_{i=1}^{n} C_i / n(n-1) = \bar{C}/(n-1) \;, \tag{4.3.3}$$

and that

$$\hat{\sigma}_J^2 = 4 \sum_{i=1}^{n} (C_i - \bar{C})^2 / (n-1)(n-2)^2 \quad , \tag{4.3.4}$$

where

$$\bar{C} = \sum_{i=1}^{n} C_i / n \quad .$$

Fligner and Rust obtained the statistic

$$K^* = n^{1/2} \hat{\tau} / \hat{\sigma}_J$$

$$= \frac{n-2}{[n(n-1)]^{1/2}} \cdot \frac{\sum_{i=1}^{n} C_i}{2[\sum_{i=1}^{n} (C_i-\bar{C})^2]^{1/2}} , \qquad (4.3.5)$$

and observed that, since $C_1, \ldots, C_n$ depend on the observations through their marginal rankings, K^* is distribution-free under the hypothesis of independence, and that the tests based on $\hat{\tau}$ and K^* have equivalent consistency classes and asymptotic relative efficiencies. They further note that an advantage of K^* over $\hat{\tau}$ is that K^* is also asymptotically distribution-free under the hypothesis H_0: $\tau=0$. One drawback to using the Fligner-Rust statistic is that $\hat{\sigma}_J^2$ may be identically zero even in non-extreme cases. In a preliminary simulation study, we have discovered several rank configurations such as the one given below where $\hat{\sigma}_J^2 = 0$. When, for example, the ranks are

Rank (Y):	1	2	3	4	5	6	7	8
Rank (Z):	5	6	7	8	1	2	3	4 ,

$C_i = -1$, for $i = 1, 2, \ldots, 8$, and therefore $\hat{\sigma}_J^2 = 0$. For extreme cases of "perfect concordance" or "perfect discordance," it is reasonable to assume a very small value for $\hat{\sigma}_J^2$ (i.e., a very large value for K^*), thereby rejecting the null hypothesis that $\tau=0$. However, such a procedure should not be used for situations similar to the one given above where $\hat{\tau} = -1/7$ and hence no indication of either concordance or discordance is present.

Another estimator of $\mathrm{Var}[\hat{\tau}]$ was proposed by Noether (1967). Using the notation developed above, his variance estimator may be written as

$$\hat{\mathrm{Var}}[\hat{\tau}] = \frac{1}{\binom{n}{2}}\left[\frac{2}{n(n-1)}\sum_{i=1}^{n}(C_i-\bar{C})^2 + \hat{\tau}^2 - 1\right].$$

A disadvantage of this variance estimator is that it may be negative. For example, for the rank configurations given above, $C_i = \bar{C} = -1$, $i = 1, 2, \ldots, n$, and $\hat{\tau} = -1/7$ so that

$$\hat{\mathrm{Var}}[\hat{\tau}] = -\frac{96}{49n(n-1)} .$$

We propose a variance estimator which is guaranteed to be positive except for the extreme cases of $\hat{\tau} = \pm 1$. This is the consistent estimator of $\mathrm{Var}[\hat{\tau}]$ based on the sample estimators of ζ_1 and ζ_2, similar to those considered by Randles, Fligner, Policello and Wolfe (1980) and previously developed by Sen (1960). The variance of $\hat{\tau}$ is given in (4.2.5) as

$$\mathrm{Var}[\hat{\tau}] = \frac{1}{\binom{n}{2}}[2(n-2)\,\zeta_1 + \zeta_2] \text{ , where}$$

$$\zeta_1 = \mathrm{Var}[h_1(\underset{\sim}{X}_1)] \quad \text{and} \quad \zeta_2 = 1 - \tau^2 ,$$

with

$$h_1(\underset{\sim}{x}) = E[h(\underset{\sim}{X}_1,\underset{\sim}{X}_2)\,|\,\underset{\sim}{X}_1 = \underset{\sim}{x}] .$$

Since $E[h_1(\underset{\sim}{X}_1)] = \tau$, and taking $\hat{\tau}$ to be the estimator of τ, the sample estimator of ζ_1 may be written as

$$\hat{\zeta}_1 = \frac{1}{n} \sum_{i=1}^{n} [\hat{h}_1(\underset{\sim}{X}_i) - \hat{\tau}]^2 \quad , \tag{4.3.6}$$

where

$$\hat{h}_1(\underset{\sim}{X}_i) = \frac{1}{(n-1)} \sum_{\substack{j=1 \\ j \neq i}}^{n} h(\underset{\sim}{X}_i, \underset{\sim}{X}_j)$$

$$= \frac{1}{(n-1)} \sum_{\substack{j=1 \\ j \neq i}}^{n} \mathrm{Sgn}\{(Y_i - Y_j)(Z_i - Z_j)\} \ . \tag{4.3.7}$$

Taking $\hat{\zeta}_2 = 1 - \hat{\tau}^2$ to be the estimator of ζ_2, our proposed estimator for the variance of $\hat{\tau}$ is

$$\hat{\mathrm{Var}}[\hat{\tau}] = \frac{1}{\binom{n}{2}} [2(n-2)\ \hat{\zeta}_1 + \hat{\zeta}_2] \ .$$

Using the notation of Fligner and Rust, we see from (4.3.2) and (4.3.3) that

$$\hat{h}(\underset{\sim}{X}_i) = \frac{1}{n-1} C_i \ , \text{ so that}$$

$$\hat{\zeta}_1 = \frac{1}{n} \sum_{i=1}^{n} [C_i/(n-1) - \bar{C}/(n-1)]^2$$

$$= \frac{1}{n(n-1)^2} \sum_{i=1}^{n} [C_i - \bar{C}]^2 ,$$

which is the same as the Fligner-Rust estimator of ζ_1 used in expression (4.3.4) with (n-1) and (n-2) being replaced by n and (n-1), respectively. It follows that

$$\hat{\mathrm{Var}}[\hat{\tau}] = \frac{1}{\binom{n}{2}} \left[\frac{2(n-2)}{n(n-1)^2} \sum_{i=1}^{n} (C_i - \bar{C})^2 + 1 - \hat{\tau}^2\right]$$

and the corresponding test statistic is

$$K^*_{RS} = \hat{\tau}/[\hat{\mathrm{Var}}(\hat{\tau})]^{1/2} . \qquad (4.3.8)$$

As with K^*, K^*_{RS} is distribution-free under the hypothesis of independence, and is asymptotically distribution-free under the more general hypothesis $\tau=0$.

The null distribution of K^*_{RS} was generated by a simulation study based on 10,000 replications. For each replication, two independent random samples each of size n were generated from the standard normal distribution using the IMSL library. At each stage, K^*_{RS} was calculated and a count was kept for each possible value up to three decimal places. The upper tail critical values (rounded to 2 decimal places) of K^*_{RS} for selected α-levels and for n = 6(1)30 are given in Table 4.1.

Table 4.1

The Null Distribution of K_{RS}^{*}

Selected values of K_{RS}^{*}: Upper tail critical values of the distribution of K_{RS}^{*} under the hypothesis of independence.

n	α=0.10	α=0.05	α=0.025	α=0.01	α=0.005
6	1.31	1.90	2.62	3.29	4.50
7	1.38	1.87	2.41	3.36	4.14
8	1.25	1.78	2.22	2.89	3.42
9	1.27	1.76	2.28	3.00	3.46
10	1.27	1.74	2.17	2.83	3.40
11	1.28	1.72	2.23	2.76	3.20
12	1.27	1.68	2.11	3.75	3.25
13	1.24	1.68	2.10	2.72	3.20
14	1.23	1.66	2.07	2.63	3.10
15	1.25	1.70	2.13	2.61	2.98
16	1.24	1.67	2.07	2.62	2.98
17	1.24	1.63	2.07	2.58	2.97
18	1.26	1.69	2.03	2.58	2.98
19	1.26	1.69	2.06	2.52	2.98
20	1.20	1.63	1.97	2.41	2.79
21	1.26	1.70	2.12	2.60	2.92
22	1.25	1.62	2.05	2.43	2.77
23	1.26	1.66	2.04	2.58	2.94
24	1.25	1.67	2.07	2.55	2.84
25	1.22	1.65	2.02	2.55	2.85
26	1.24	1.63	2.01	2.43	2.77
27	1.25	1.66	1.98	2.39	2.71
28	1.25	1.63	2.00	2.47	2.81
29	1.27	1.68	2.04	2.54	2.84
30	1.23	1.62	2.00	2.44	2.77

4.4 Empirical Power Comparisons

The performances of three statistics based on Kendall's $\hat{\tau}$ and four statistics based on Pearson's R were investigated through a Monte Carlo simulation study, each with 1000 replications. The statistics considered were the following

1) $K = \binom{n}{2} \hat{\tau}$ was compared to table A.21 of Hollander and Wolfe (1973).

2) $ZK = K/[n(n-1)(2n+5)/18]^{1/2}$, which is Kendall's statistic standardized by the variance under the hypothesis of independence, was compared to the 0.05 upper critical value of the standard normal distribution $Z_{0.05} = 1.645$. For n=8, both a correction for continuity (adjusted by 1 rather than by 0.5 since K takes on only even values) and randomization for an exact α=0.05 level through the use of a Uniform [0,1] random variable were employed.

3) Our proposed statistic, K^{*}_{RS}, was compared to the simulated critical values given in table 4.1.

4) $T = R \left\{ \frac{(n-2)}{1-R^2} \right\}^{1/2}$ was compared to the upper 0.05 cut-off value of the Student's t-distribution with (n-2) degrees of freedom.

5) The standardized Fisher's Z,

$$FZ = Z/\left[\frac{1}{n-3} \right]^{1/2},$$

where

$$Z = \frac{1}{2} \ln \left[(1+R)/(1-R)\right]$$

was compared to $Z_{0.05} = 1.645$.

6) and 7)

$$RJ = \frac{R}{[Var_J(R)]^{1/2}}$$

and

$$ZJ = \frac{Z}{[Var_J(Z)]^{1/2}} \quad ,$$

where $Var_J(R)$ and $Var_J(Z)$ are the jackknife estimators of the variances of R and Z, respectively, were compared to $Z_{0.05} = 1.645$. The use of these jackknife estimators was motivated by our suspicion that they may improve the performances of the tests based on R or Z when the assumptions of normality and/or independence were no longer present. Fisher's Z transform was included in this study not only for completeness but also because of some of its desirable properties such as its stabilized variance, and the fact that it is "more nearly normal" than R. Furthermore, most advocates of the jackknife recommend variance stabilizing transformations to "keep the jackknife on scale and thus prevent distortion of the results" (see, for example, Hinkley, 1977, 1978, and Miller, 1974). The jackknife estimators of R and Z were obtained by a procedure similar to that given in Hinkley (1978). First, we calculate the pseudovalues

$$PR^{(i)} = nR - (n-1)R^{(i)}$$

and

$$PZ^{(i)} = nZ - (n-1)Z^{(i)} \;, \; i=1,2,\ldots,n$$

where $R^{(i)}$ is the product moment correlation coefficient based on a sample of size (n-1) obtained by deleting the i^{th} pair, and $Z^{(i)}$ is the corresponding Fisher transform; i.e.,

$$Z^{(i)} = \frac{1}{2} \ln \left[\frac{1 + R^{(i)}}{1 - R^{(i)}}\right] \;.$$

The sample variances of the pseudovalues are then given by

$$VR = \frac{1}{(n-1)} \sum_{i=1}^{n} [PR^{(i)} - \overline{PR}]^2$$

and

$$VZ = \frac{1}{(n-1)} \sum_{i=1}^{n} [PZ^{(i)} - \overline{PZ}]^2 \;,$$

where

$$\overline{PR} = \frac{1}{n} \sum_{i=1}^{n} PR^{(i)} \qquad \text{and} \qquad \overline{PZ} = \frac{1}{n} \sum_{i=1}^{n} PZ^{(i)} \;.$$

The recommended variance estimates of R and Z are then

$$Var_J(R) = \frac{VR}{n} \qquad \text{and} \qquad Var_J(Z) = \frac{VZ}{n} \;.$$

In the computer algorithm to calculate these jackknife estimators, some useful recursive relations were used which enable one to update sample variances and covariances when the sample is augmented by an additional observation. These relations are derived to be

$$\sum_{i=1}^{n} (X_i-\bar{X}_n)^2 = \sum_{i=1}^{n-1} (X_i-\bar{X}_{n-1})^2 + \frac{n-1}{n} (X_n-\bar{X}_{n-1})^2$$

and

$$\sum_{i=1}^{n} (X_i-\bar{X}_n)(Y_i-\bar{Y}_n) = \sum_{i=1}^{n-1} (X_i-\bar{X}_{n-1})(Y_i-\bar{Y}_{n-1}) + \frac{n-1}{n}(X_n-\bar{X}_{n-1})(Y_n-\bar{Y}_{n-1})$$

where $$\bar{a}_k = \frac{1}{k} \sum_{i=1}^{k} a_i \ .$$

The results of the Monte Carlo study based on 1000 samples each comprised of n=8 and n=20 pairs of observations are given in tables 4.2-4.5. The empirical sizes (powers) corresponding to the seven tests listed above were computed for several bivariate distributions. For the hypothesis of independence, the pairs (Y,Z) were formed by letting

$$Y = W_1 + \Delta W_3$$

and (4.4.1)

$$Z = W_2 + \Delta W_3$$

where W_1, W_2 and W_3 are independent random variables, so that the hypothesis of independence is equivalent to testing $\Delta=0$. For each of the 1000 iterations, 3n i.i.d. random variates were generated from a specific distribution using IMSL subroutines. These were divided into three groups, each of size n (n=8 and n=20), to obtain $\{W_{1i}\}$, $\{W_{2i}\}$ and $\{W_{3i}\}$, i = 1, 2, . . . , n. The pairs (Y_i,Z_i), i = 1, 2, . . . , n, were then obtained by relations (4.4.1), for Δ = 0.0, 1.0 and 2.0 (i.e., when Corr(Y,Z) = 0.0, 0.5 and 0.8, respectively). The seven statistics mentioned earlier were calculated from these pairs, and were compared to their corresponding cut-off values to obtain the empirical powers. The results for the standard normal, the Uniform [0,1], and the Cauchy distributions are given in table 4.3.

To test the hypothesis H_0: $\tau=0$, the seven statistics under investigation were calculated on (Y_1,Z_1), . . . , (Y_n,Z_n), n=8 and n=20, but here the (Y,Z) pairs were generated from such bivariate distributions as the bivariate Cauchy, the Pearson Type II and the Pearson Type VII distributions (see Johnson and Ramberg, 1977). In the case of such elliptically symmetric distributions, $\lambda=0$ is equivalent to $\rho\equiv$Corr(Y,Z)=0 which in turn is equivalent to $\tau=0$. To generate these bivariate observations, we have adopted the procedures given by Johnson and Ramberg (1977). To form a (Y,Z) pair, we first implement IMSL subroutines to obtain two random independent U[0,1] variates, U_1 and U_2. For each of the three bivariate distributions mentioned above, U_1 and U_2 are transformed into two uncorrelated variables X_1 and X_2, by appropriate transformations discussed below. The pair (Y,Z) is then obtained by

$$Y = X_1$$

and (4.4.2)

$$Z = \lambda X_1 + (1-\lambda^2)^{1/2} X_2$$

where $0 \leq \lambda \leq 1$, and $Corr(Y,Z) = \lambda$, if X_1 and X_2 have finite equal variances. For the bivariate Cauchy distributions which is a heavy-tailed distribution with no moments, the transforms X_1 and X_2 are obtained as follows.

$$X_1 = (U_1^{-2} - 1)^{1/2} \, Cos(2\pi U_2)$$

and

$$X_2 = (U_1^{-2} - 1)^{1/2} \, Sin(2\pi U_2) \quad .$$

The Pearson Type II is a light-tailed distribution which converges to the bivariate normal distribution as the shape parameter ν increases to infinity. Here X_1 and X_2 are obtained by

$$X_1 = (1 - U_1^{1/\nu})^{1/2} \, Cos(2\pi U_2)$$

and

$$X_2 = (1 - U_1^{1/\nu})^{1/2} \, Sin(2\pi U_2) \quad .$$

The Pearson Type VII distribution is more heavy-tailed than the bivariate normal distribution, with the tail weight increasing as the parameter ν decreases. X_1 and X_2 are given by

$$X_1 = (U_1^{1/(1-\nu)} - 1)^{1/2} \, Cos(2\pi U_2)$$

and

Table 4.2

Relative Frequency of Rejecting H_0

(nominal α=0.05)

Distribution of W_1,W_2,W_3	Δ	Tests Based on $\hat{\tau}$ K	ZK	K^*_{RS}	Tests Based on R T	FZ	RJ	ZJ
				n=8				
Normal	0.0	0.048	0.048	0.055	0.042	0.040	0.087	0.060
	1.0	0.311	0.311	0.314	0.376	0.370	0.487	0.330
	2.0	0.731	0.731	0.737	0.853	0.850	0.882	0.752
Uniform	0.0	0.048	0.048	0.055	0.051	0.049	0.090	0.057
	1.0	0.288	0.288	0.306	0.347	0.340	0.490	0.343
	2.0	0.754	0.754	0.772	0.894	0.891	0.933	0.838
Cauchy	0.0	0.046	0.046	0.061	0.058	0.057	0.076	0.048
	1.0	0.340	0.340	0.331	0.409	0.407	0.384	0.204
	2.0	0.529	0.529	0.509	0.572	0.568	0.548	0.365
				n=20				
Normal	0.0	0.060	0.059	0.059	0.056	0.056	0.077	0.065
	1.0	0.688	0.684	0.676	0.766	0.762	0.777	0.724
	2.0	0.994	0.994	0.992	0.998	0.998	0.997	0.994
Uniform	0.0	0.060	0.059	0.059	0.066	0.065	0.072	0.058
	1.0	0.703	0.699	0.721	0.779	0.776	0.857	0.817
	2.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0
Cauchy	0.0	0.047	0.046	0.047	0.057	0.056	0.028	0.024
	1.0	0.622	0.619	0.568	0.512	0.512	0.344	0.227
	2.0	0.902	0.901	0.850	0.717	0.717	0.541	0.356

Table 4.3

Relative Frequency of Rejecting H_0

(nominal α=0.05)

Distn. of (Y,Z)	ν	λ	Tests Based on $\hat{\tau}$: K	ZK	K^*_{RS}	Tests Based on R: T	FZ	RJ	ZJ
					n=8				
	1.0	0.0	0.025	0.025	0.033	0.030	0.028	0.094	0.058
P		0.5	0.283	0.285	0.330	0.334	0.325	0.530	0.383
		0.8	0.770	0.772	0.770	0.877	0.871	0.929	0.833
E									
	5.0	0.0	0.039	0.039	0.051	0.052	0.049	0.107	0.057
A		0.5	0.303	0.305	0.315	0.355	0.350	0.497	0.369
		0.8	0.770	0.772	0.759	0.860	0.857	0.904	0.775
R									
					n=20				
S									
	1.0	0.0	0.023	0.023	0.036	0.025	0.025	0.059	0.05
O		0.5	0.743	0.740	0.791	0.810	0.807	0.882	0.865
		0.8	1.0	1.0	1.0	1.0	1.0	1.0	1.0
N									
	5.0	0.0	0.042	0.041	0.049	0.035	0.034	0.061	0.053
		0.5	0.711	0.708	0.708	0.778	0.777	0.830	0.782
II		0.8	0.999	0.999	0.998	1.0	1.0	1.0	1.0

Table 4.4

Relative Frequency of Rejecting H_0

(nominal α=0.05)

Distn. of (Y,Z)	ν	λ	Tests Based on $\hat{\tau}$			Tests Based on R			
			K	ZK	K^*_{RS}	T	FZ	RJ	ZJ
					n=8				
PEARSON VII	2.0	0.0	0.076	0.077	0.075	0.143	0.139	0.130	0.075
		0.5	0.354	0.356	0.331	0.457	0.453	0.470	0.317
		0.8	0.732	0.734	0.707	0.793	0.789	0.802	0.630
	1.25	0.0	0.106	0.107	0.091	0.358	0.356	0.177	0.083
		0.5	0.383	0.385	0.338	0.580	0.574	0.426	0.198
		0.8	0.702	0.704	0.641	0.783	0.782	0.690	0.380
					n=20				
	2.0	0.0	0.078	0.077	0.059	0.214	0.214	0.104	0.073
		0.5	0.670	0.669	0.593	0.672	0.671	0.583	0.473
		0.8	0.985	0.985	0.968	0.945	0.944	0.905	0.813
	1.25	0.0	0.101	0.100	0.070	0.434	0.433	0.166	0.061
		0.5	0.643	0.642	0.511	0.659	0.659	0.444	0.198
		0.8	0.962	0.962	0.915	0.836	0.835	0.703	0.356

Table 4.5

Relative Frequency of Rejecting H_0

(nominal α=0.05)

Distn. of (Y,Z)	λ	Tests Based on $\hat{\tau}$			Tests Based on R			
		K	ZK	K^*_{RS}	T	FZ	RJ	ZJ
				n=8				
BIVARIATE CAUCHY	0.0	0.089	0.090	0.084	0.245	0.243	0.151	0.091
	0.5	0.368	0.370	0.323	0.516	0.515	0.443	0.263
	0.8	0.727	0.729	0.683	0.782	0.776	0.744	0.522
				n=20				
	0.0	0.091	0.090	0.065	0.350	0.350	0.131	0.075
	0.5	0.653	0.649	0.551	0.660	0.660	0.481	0.326
	0.8	0.977	0.977	0.947	0.878	0.878	0.793	0.599

$$X_2 = (U_1^{1/(1-\nu)} - 1)^{1/2} \text{Sin}(2\pi U_2) \quad .$$

Note that for $\nu = 1.5$, the Pearson VII is equivalent to the bivariate Cauchy distribution. The results of the Monte Carlo study for $\lambda = 0.0$, 0.5 and 0.8 and for selected values of ν are given in tables 4.3-4.5.

4.5 Conclusions and Recommendations

In many cases, it is difficult to compare the powers of these tests especially when the corresponding α-levels are highly different for the different tests. In this discussion we present what we believe to be a reasonable set of conclusions drawn from our study. One such conclusion is that for the hypothesis of independence, the tests based on R, namely T and FZ are hignly robust in the sense of having stable sizes and powers, as may be seen in table 4.2. This was to be expected for light-tailed distributions, as was indicated by the ARE calculations given in section 1.2. However, for n=20 the tests based on Kendall's tau have slightly higher powers for a heavy-tailed distribution such as the Cauchy, although for n=8 the performance of the tests T and FZ is comparable to, if not better than, that of the tests based on $\hat{\tau}$. The tests based on R also do well for the hypothesis H_0: $\tau=0$ when the observations come from a light-tailed bivariate distribution such as the Pearson II (see table 4.3). For n=8, both T and FZ perform remarkably well in terms of holding their α-levels and powers, while for n=20, the tests using the jackknife variance estimators, i.e., RJ and ZJ, do considerably better, followed by the test K^*_{RS}. For the hypothesis H_0: τ=0, and for heavier-tailed bivariate distributions such as the Pearson VII or the bivariate Cauchy the tests based on Kendall's tau do extremely well. Except for the test ZJ, which is Fisher's Z transform standardized by the jackknife estimator of standard error, all tests based on R have highly inflated α-levels, and hence should not be included in any power comparisons. Of the remaining tests, those based on $\hat{\tau}$ exhibit the highest empirical powers. In particular, our test K^*_{RS} performs the best both in terms of empirical α-level and power.

In summary, we note that for the hypothesis of independence the tests based on Pearson's R are, in most cases, remarkably robust in terms of both size and power. For the hypothesis τ=0, the tests based on K^*_{RS} are consistently better except in the case of the Pearson II distribution where ZJ has slightly higher powers. However, in practice one must take into consideration the ease with which a particular statistic is calculated. As can be seen from the previous section, the computation of a statistic such as ZJ is very tedious compared to that of K^*_{RS} which is a function of the C_i's which are naturally calculated in a Kendall's tau problem. Based on the above discussions, our final recommendations are

1) For the hypothesis of independence, we recommend using a simple test based on R such as T or FZ, except for large n ($\geq$20) and heavy-tailed distributions where we recommend the use of a test based on the ordinary Kendall's tau such as K or ZK.

2) For the hypothesis H_0: $\tau=0$, we recommend a test based on our statistic K_{RS}^* in all situations. Furthermore, it is important to note that K_{RS}^* may also be used to construct confidence intervals for τ. For small samples we recommend the use of table 4.1, while for large samples (n>30) one may use the appropriate percentiles of the standard normal distribution.

CHAPTER FIVE
MONTE CARLO RESULTS AND CONCLUSIONS

5.1 Introduction

In Chapters 2 and 3, we discussed two tests for partial correlation based respectively on T_n, Kendall's tau statistic calculated on the residuals, and R_n, Pearson's partial correlation coefficient. Based on the values of ARE(T_n,R_n) calculated in Chapter 3, we concluded that, for large samples and under the null hypothesis of the independence of E and E' (the "error variables" in the linear models relating Y to X, and Z to X, respectively), T_n performs better than R_n for heavy tailed distributions. In Chapter 4, we studied the usual correlation problem and discussed several statistics for testing the null hypothesis H_0: $\tau=0$, where τ was Kendall's correlation coefficient between the variables Y and Z. In this chapter, a Monte Carlo study is used to investigate the performances of the tests based on T_n, R_n, and statistics similar to those discussed in Chapter 4 but here calculated on the residuals from the fit involving the covariate X.

In section 5.2, we shall discuss statistics similar to the ones studied in Chapter 4 but modified to fit the partial correlation setting, and tabulate their simulated null distributions. Section 5.3 contains a description of our Monte Carlo study and the tables of results. Section 5.4 contains our overall conclusions and recommendations. In section 5.5, we give a brief list of related topics open for future research and investigation.

5.2 More Tests for Partial Correlation

In this section, we shall develop some statistics for testing a broader null hypothesis than that of the independence between E and E'. In particular, we shall be interested in testing

$$H_o: \ \tau = 0 \qquad \text{versus} \qquad H_a: \ \tau > 0 , \tag{5.2.1}$$

where

$$\tau = P\{(E_1-E_2)(E_1'-E_2')>0\} - P\{(E_1-E_2)(E_1'-E_2')<0\} . \tag{5.2.2}$$

The two primary measures for partial correlation discussed in Chapters 2 and 3 are

$$T_n = \frac{1}{\binom{n}{2}} \sum_{i<j} \mathrm{Sgn}\{(U_i-U_j)(V_i-V_j)\} \tag{5.2.3}$$

and

$$R_n = \frac{\sum_{i=1}^{n} (U_i-\bar{U})(V_i-\bar{V})}{\{\sum_{i=1}^{n} (U_i-\bar{U})^2 \sum_{i=1}^{n} (V_i-\bar{V})^2\}^{1/2}} , \tag{5.2.4}$$

where $(U_1,V_1), \ldots, (U_n,V_n)$ are the residuals obtained from fitting the linear models

$$Y_i = \alpha_1 + \beta_1 X_i + E_i$$

and (5.2.5)

$$Z_i = \alpha_2 + \beta_2 X_i + E_i' \quad , \; i=1,2,\ldots,n.$$

To test the hypotheses of (5.2.1), in addition to the statistics T_n and R_n, we also use statistics similar to those discussed in Chapter 4. One such statistic is K_n^* which is the statistic K_{RS}^* given in (4.3.8) but applied to the residuals. That is, if

$$C_i = \sum_{\substack{j=1 \\ j\neq i}}^{n} \text{Sgn}\{(U_i-U_j)(V_i-V_j)\} ,$$

the statistic K_n^* may be written as

$$K_n^* = \frac{T_n}{\binom{n}{2}^{1/2} \left\{\frac{2(n-2)}{n(n-1)^2} \sum_{i=1}^{n} (C_i-\bar{C})^2 + 1 - T_n^{\,2}\right\}^{1/2}} \tag{5.2.6}$$

where $\bar{C} = \frac{1}{n}\sum_{i=1}^{n} C_i$, and T_n is Kendall's tau applied to the residuals as given in (5.2.2).

The distribution of K_n^* under the null hypothesis that $\tau=0$ was generated by a Monte Carlo simulation study in two cases: (i) when the residuals were obtained by the OLS fit and (ii) when they were obtained by the LAV fit. In each of these two cases, the residuals were obtained from the models

$$Y_i = X_i + E_i$$

and

$$Z_i = X_i + E_i^{'} \quad , \; i=1,2,\ldots,n \; ,$$

where X_i, $i = 1,2,\ldots,n$, are i.i.d. standard normal variables generated by IMSL subroutines, and $(E_i,E_i^{'})$, $i = 1,2,\ldots,n$, are pairs of observations from the Pearson Type VII distribution with λ=0 (i.e., τ=0) and ν=2 and generated by the procedures described in section 4.4. From these residuals the statistic K_n^* was calculated and its value recorded. This process was repeated 10,000 times. The upper tail critical values of K_n^* for selected values of α and for n = 6(1)20 are given in Table 5.1 (the OLS fit) and Table 5.2 (the LAV fit).

It must be noted that the use of the Pearson Type VII distribution with ν=2 to generate the null (τ=0) distribution of K_n^* was not altogether arbitrary. This choice was motivated by the fact that this particular distribution is "close" to the bivariate normal distribution in terms of having moments and in terms of tail weight (it has a slightly heavier tail than the bivariate normal distribution), but it is more appropriate than the bivariate normal distribution for testing the null hypothesis τ=0 since under the Pearson Type VII distribution, λ=0 (τ=0 and ρ=0) does not necessarily imply that E and $E^{'}$ are statistically independent.

Table 5.1

The Null Distribution of K_n^* (OLS fit)

n	α=0.10	α=0.05	α=0.025	α=0.01	α=0.005
6	1.90	2.62	-	-	-
7	1.87	2.60	3.36	4.72	6.48
8	1.80	2.44	3.15	4.61	5.22
9	1.68	2.35	3.05	4.11	5.00
10	1.66	2.26	2.96	4.14	4.95
11	1.64	2.24	2.89	4.05	4.96
12	1.62	2.21	2.89	3.88	4.95
13	1.59	2.18	2.83	3.74	4.65
14	1.57	2.18	2.83	3.72	4.5
15	1.56	2.12	2.77	3.52	4.41
16	1.54	2.11	2.70	3.57	4.45
17	1.52	2.08	2.71	3.56	4.30
18	1.51	2.07	2.65	3.43	4.14
19	1.50	2.01	2.63	3.38	3.98
20	1.52	2.03	2.62	3.44	4.00

Table 5.2

The Null Distribution of K_n^* (LAV fit)

n	α=0.10	α=0.05	α=0.025	α=0.01	α=0.005
6	1.67	2.90	-	-	-
7	1.59	2.40	3.36	4.72	6.45
8	1.43	2.05	2.80	3.94	5.22
9	1.44	2.00	2.61	3.60	4.44
10	1.39	1.96	2.52	3.30	4.18
11	1.42	1.97	2.48	3.15	3.71
12	1.41	1.92	2.38	3.07	3.73
13	1.38	1.88	2.36	3.00	3.41
14	1.37	1.85	2.28	3.01	3.48
15	1.34	1.82	2.29	2.98	3.55
16	1.36	1.81	2.30	2.89	3.26
17	1.35	1.78	2.22	2.87	3.31
18	1.33	1.75	2.21	2.72	3.18
19	1.34	1.76	2.16	2.78	3.15
20	1.34	1.76	2.21	2.76	3.17

5.3 The Monte Carlo Study

In this section we compare the performances of the tests based on T_n, R_n and K_n^* through a Monte Carlo simulation study with 1000 replications. The hypotheses of interest are

$$H_0: \text{ E and } E' \text{ are independent} \tag{5.3.1}$$

and

$$H_0^*: \tau = 0 ,$$

where τ is as defined in (5.2.2), and E and E' are the error variables of the model structures (5.2.5). Throughout tnis study we have taken the variable X to have the standard normal distribution, and have let the pair (E,E') assume a variety of different bivariate distributions.

For the hypothesis of independence, the class of alternatives is defined by the "trivariate reduction" model given by

$$E = W_1 + \Delta W_3$$

and

$$E' = W_2 + \Delta W_3 ,$$

where W_1, W_2 and W_3 are mutually independent continuous random variables, and Δ is a constant. The hypothesis of independence (5.3.1) is then equivalent to

$$H_0: \Delta = 0 .$$

For the hypothesis (5.3.2), we have taken the variable (E,E') to have an elliptically symmetric distribution with "association parameter" λ, so that the hypothesis (5.3.2) is equivalent to

$$H_o: \lambda = 0 ,$$

where λ is as given in (4.4.2).

The variables $X_1, X_2, \ldots, X_n$ were generated by the IMSL subroutine GGNML, and the pairs $(E,E',), \ldots, (E_n,E_n')$ were obtained by the exact same procedures used to generate the variables $(Y_1,Z_1), \ldots, (Y_n,Z_n)$ in section 4.4. The variables of interest under the partial correlation setting were then formed by calculating

$$Y_i = X_i + E_i$$

and

$$Z_i = X_i + E_i' , \quad i=1,2,\ldots,n .$$

From the above linear models, pairs of residuals $(U_1,V_1), \ldots, (U_n,V_n)$ were obtained from (i) the OLS fit, and (ii) the LAV fit, and from each of the two sets of residual pairs the statistics T_n, K_n^* and R_n were calculated. Based on these statistics, the performances of the following seven tests were compared.

Tests based on T_n:

(i) $T_1 = \binom{n}{2}T_n$ was compared to table A.21 of Hollander and Wolfe (1973).

(ii) $T_2 = \binom{n}{2}T_n$ was compared to the tables of the simulated distributions of T_n under the hypothesis of independence. It was compared to values of table 2.1 for the OLS fit, and table 2.2 for the LAV fit.

(iii) $$T_3 = \frac{T_n}{\{(2n+5)/18n(n-1)\}^{1/2}},$$

which is T_n standardized by the variance of the ordinary Kendall's tau under independence, was compared to the upper $\alpha=0.05$ critical value of the standard normal distribution $Z_{0.05} = 1.645$.

For each of the above three tests randomization was employed to obtain an exact $\alpha=0.05$ level.

<u>Tests based on K_n^*:</u>

The three tests K_1^*, K_2^* and K_3^*, respectively, were obtained by comparing K_n^* to

(i) the $\alpha=0.05$ cutoff values of the distribution of K_{RS}^* (K_n^* under the ordinary correlation problem given in table 4.1),

(ii) the $\alpha=0.05$ cutoff values of the simulated null distribution of K_n^* when (E,E') has the bivariate normal distribution. Only selected cutoff values were generated for completion. For reasons we discussed in the previous section, we recommend using tables 5.1 and 5.2 which contain the null distribution of K_n^* when (E,E') has the Pearson VII distribution, and

(iii) by comparing K_n^* to the $\alpha=0.05$ critical value of table 5.1 (for the OLS fit) and table 5.2 (for the LAV fit).

Tests based on R_n:

(i) $R_1 = R_n \left\{ \frac{n-3}{1-R^2} \right\}^{1/2}$ was compared to the upper 0.05 cutoff value of the Student's t-distribution with (n-3) degrees of freedom.

(ii) $R_2 = \frac{Z_n}{\left\{\frac{1}{n-3}\right\}^{1/2}}$ where $Z_n = \frac{1}{2} \ln \left\{\frac{1+R_n}{1-R_n}\right\}$ (5.3.3)

was compared to $Z_{0.05} = 1.645$.

(iii) $R_3 = R_n / \{Var_J(R_n)\}^{1/2}$ was compared to $Z_{0.05} = 1.645$.

(iv) $R_4 = Z / \{Var_J(Z_n)\}^{1/2}$ was compared to $Z_{0.05} = 1.645$, where Z is as given in (5.3.3).

The jackknife variance estimators $Var_J(R_n)$ and $Var_J(Z_n)$ were obtained by the procedures discussed in section 4.4 but applied here to the residual pairs.

The relative frequencies of rejecting H_0 for sample sizes n=8 and n=20, and for various distributions are given in tables 5.3-5.12. Tables 5.3-5.6 contain the results for the hypothesis of conditional independence where the class of alternatives is given by the "trivariate reduction" model. Tables 5.7-5.12 contain the results when (E,E') has an elliptically symmetric bivariate distribution.

Table 5.3

Relative Frequency of Rejecting H_0 (OLS fit)
Trivariate Reduction Model

(nominal α=0.05)

Distribution of W_1, W_2, W_3	Δ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 8						
Normal	0.0	0.061	0.049	0.062	0.069	0.054	0.034	0.051	0.073	0.112	0.077
	1.0	0.322	0.278	0.324	0.339	0.281	0.184	0.331	0.385	0.491	0.349
	1.7	0.629	0.565	0.631	0.628	0.577	0.443	0.679	0.728	0.788	0.635
	2.0	0.721	0.664	0.723	0.715	0.664	0.557	0.774	0.822	0.866	0.733
Uniform	0.0	0.048	0.042	0.048	0.055	0.046	0.028	0.045	0.068	0.112	0.065
	1.0	0.326	0.281	0.328	0.333	0.273	0.192	0.306	0.371	0.502	0.339
	1.7	0.642	0.587	0.644	0.649	0.591	0.464	0.723	0.772	0.842	0.692
	2.0	0.759	0.705	0.761	0.767	0.715	0.568	0.824	0.868	0.906	0.799
Cauchy	0.0	0.047	0.041	0.047	0.059	0.044	0.024	0.049	0.066	0.083	0.052
	1.0	0.366	0.340	0.368	0.361	0.318	0.257	0.364	0.391	0.442	0.313
	1.7	0.477	0.452	0.479	0.479	0.441	0.381	0.501	0.523	0.560	0.427
	2.0	0.525	0.500	0.526	0.523	0.487	0.416	0.545	0.572	0.596	0.461

Table 5.4

Relative Frequency of Rejecting H_o (LAV fit)
Trivariate Reduction Model

(nominal α=0.05)

Distribution of W_1, W_2, W_3	Δ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 8						
Normal	0.0	0.043	0.038	0.044	0.049	0.041	0.037	0.045	0.062	0.091	0.054
	1.0	0.247	0.236	0.248	0.282	0.254	0.227	0.280	0.337	0.421	0.288
	1.7	0.524	0.507	0.526	0.543	0.500	0.465	0.607	0.663	0.725	0.552
	2.0	0.622	0.599	0.624	0.647	0.603	0.573	0.713	0.759	0.805	0.645
Uniform	0.0	0.039	0.039	0.039	0.053	0.047	0.039	0.036	0.05	0.094	0.05
	1.0	0.266	0.249	0.268	0.295	0.253	0.229	0.260	0.328	0.444	0.288
	1.7	0.565	0.538	0.567	0.580	0.530	0.503	0.641	0.707	0.777	0.618
	2.0	0.662	0.633	0.664	0.676	0.632	0.602	0.752	0.803	0.861	0.716
Cauchy	0.0	0.040	0.037	0.040	0.053	0.044	0.037	0.040	0.055	0.062	0.037
	1.0	0.289	0.277	0.291	0.297	0.269	0.245	0.342	0.369	0.330	0.201
	1.7	0.400	0.387	0.402	0.423	0.388	0.360	0.478	0.506	0.451	0.287
	2.0	0.458	0.445	0.460	0.467	0.432	0.411	0.523	0.544	0.503	0.312

Table 5.5

Relative Frequency of Rejecting H_0 (OLS fit)
Trivariate Reduction Model

(nominal α=0.05)

Distribution of W_1, W_2, W_3	Δ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 20						
Normal	0.0	0.058	0.055	0.057	0.066	0.052	0.032	0.061	0.071	0.082	0.070
	1.0	0.710	0.691	0.708	0.697	0.668	0.549	0.746	0.765	0.792	0.730
	1.7	0.985	0.983	0.985	0.977	0.970	0.946	0.990	0.992	0.990	0.987
	2.0	0.995	0.994	0.994	0.992	0.990	0.983	1.0	1.0	0.999	0.995
Uniform	0.0	0.056	0.052	0.054	0.060	0.048	0.030	0.059	0.063	0.075	0.051
	1.0	0.712	0.688	0.709	0.731	0.698	0.587	0.767	0.785	0.851	0.814
	1.7	0.994	0.993	0.993	0.993	0.993	0.983	0.998	0.998	0.998	0.998
	2.0	1.0	0.999	0.999	0.999	0.999	0.997	1.0	1.0	1.0	1.0
Cauchy	0.0	0.074	0.067	0.073	0.071	0.065	0.034	0.051	0.052	0.057	0.055
	1.0	0.598	0.586	0.596	0.560	0.544	0.458	0.524	0.532	0.429	0.284
	1.7	0.786	0.778	0.784	0.746	0.730	0.649	0.660	0.668	0.571	0.411
	2.0	0.816	0.806	0.814	0.784	0.772	0.703	0.706	0.707	0.608	0.448

Table 5.6

Relative Frequency of Rejecting H_o (LAV fit)
Trivariate Reduction Model

(nominal α=0.05)

Distribution of W_1, W_2, W_3	Δ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 20						
Normal	0.0	0.056	0.053	0.056	0.064	0.061	0.045	0.055	0.062	0.077	0.062
	1.0	0.674	0.649	0.672	0.664	0.654	0.617	0.725	0.742	0.762	0.705
	1.7	0.975	0.968	0.973	0.969	0.968	0.958	0.987	0.989	0.987	0.980
	2.0	0.992	0.992	0.992	0.987	0.987	0.983	0.998	0.999	0.999	0.993
Uniform	0.0	0.055	0.049	0.053	0.063	0.056	0.046	0.050	0.053	0.067	0.046
	1.0	0.675	0.652	0.674	0.690	0.680	0.638	0.737	0.758	0.832	0.787
	1.7	0.992	0.990	0.992	0.992	0.990	0.986	0.996	0.997	0.998	0.997
	2.0	0.999	0.999	0.999	0.999	0.999	0.997	0.999	1.0	1.0	1.0
Cauchy	0.0	0.064	0.061	0.063	0.070	0.065	0.053	0.049	0.050	0.051	0.043
	1.0	0.647	0.629	0.646	0.583	0.566	0.518	0.523	0.531	0.348	0.210
	1.7	0.842	0.831	0.841	0.798	0.787	0.760	0.662	0.666	0.512	0.322
	2.0	0.887	0.882	0.884	0.839	0.836	0.808	0.703	0.712	0.549	0.355

Table 5.7

Relative Frequency of Rejecting H_0 (OLS fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	ν	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
						n = 8						
P	1.0	0.0	0.039	0.033	0.039	0.053	0.036	0.019	0.040	0.056	0.117	0.067
		0.5	0.308	0.258	0.310	0.331	0.264	0.170	0.291	0.353	0.495	0.363
E		0.8	0.733	0.676	0.735	0.726	0.663	0.549	0.794	0.839	0.901	0.800
A	5.0	0.0	0.057	0.048	0.058	0.061	0.047	0.026	0.055	0.075	0.129	0.073
		0.5	0.315	0.275	0.317	0.323	0.263	0.190	0.321	0.377	0.490	0.369
R		0.8	0.711	0.662	0.713	0.711	0.657	0.534	0.786	0.830	0.883	0.770
S						n = 20						
O	1.0	0.0	0.031	0.027	0.030	0.050	0.037	0.021	0.028	0.030	0.062	0.050
		0.5	0.683	0.664	0.681	0.742	0.713	0.587	0.758	0.771	0.858	0.825
N		0.8	0.998	0.998	0.998	0.998	0.996	0.990	0.999	0.999	1.0	1.0
	5.0	0.0	0.052	0.046	0.050	0.056	0.048	0.029	0.044	0.049	0.077	0.062
II		0.5	0.664	0.646	0.663	0.674	0.649	0.526	0.733	0.750	0.810	0.743
		0.8	0.996	0.995	0.996	0.993	0.993	0.988	0.999	0.999	0.999	0.999

Table 5.8

Relative Frequency of Rejecting H_o (LAV fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	ν	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
						n = 8						
P	1.0	0.0	0.041	0.037	0.041	0.048	0.040	0.035	0.026	0.038	0.091	0.048
		0.5	0.239	0.225	0.241	0.281	0.250	0.221	0.244	0.291	0.429	0.303
E		0.8	0.612	0.595	0.614	0.645	0.604	0.568	0.724	0.779	0.837	0.696
A	5.0	0.0	0.046	0.042	0.046	0.056	0.047	0.040	0.041	0.059	0.099	0.059
		0.5	0.249	0.236	0.251	0.289	0.258	0.221	0.255	0.321	0.426	0.304
R		0.8	0.622	0.604	0.624	0.654	0.608	0.564	0.730	0.764	0.831	0.665
S						n = 20						
O	1.0	0.0	0.029	0.027	0.029	0.038	0.034	0.028	0.027	0.030	0.057	0.050
		0.5	0.658	0.638	0.657	0.707	0.695	0.656	0.719	0.738	0.827	0.799
N		0.8	0.996	0.996	0.996	0.996	0.996	0.996	0.998	0.998	1.0	1.0
	5.0	0.0	0.048	0.042	0.046	0.051	0.050	0.045	0.040	0.044	0.072	0.061
II		0.5	0.646	0.622	0.643	0.653	0.639	0.595	0.710	0.726	0.785	0.729
		0.8	0.994	0.994	0.994	0.993	0.991	0.990	0.998	0.998	0.999	0.998

Table 5.9

Relative Frequency of Rejecting H_0 (OLS fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	ν	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
						n = 8						
P	2.0	0.0	0.098	0.091	0.099	0.10	0.084	0.053	0.138	0.156	0.157	0.115
		0.5	0.350	0.323	0.352	0.351	0.311	0.236	0.427	0.463	0.494	0.367
E		0.8	0.719	0.677	0.721	0.701	0.641	0.551	0.760	0.799	0.820	0.691
A	1.25	0.0	0.103	0.099	0.103	0.102	0.092	0.084	0.338	0.356	0.173	0.083
		0.5	0.329	0.320	0.330	0.317	0.296	0.273	0.552	0.563	0.406	0.195
R		0.8	0.621	0.609	0.622	0.606	0.574	0.553	0.760	0.777	0.672	0.372
S						n = 20						
O	2.0	0.0	0.100	0.091	0.099	0.080	0.071	0.052	0.200	0.206	0.115	0.087
		0.5	0.653	0.639	0.652	0.606	0.579	0.473	0.662	0.668	0.588	0.499
N		0.8	0.979	0.973	0.977	0.956	0.951	0.933	0.946	0.951	0.909	0.837
	1.25	0.0	0.390	0.387	0.389	0.359	0.357	0.335	0.415	0.420	0.294	0.171
VII		0.5	0.668	0.658	0.665	0.635	0.624	0.586	0.654	0.658	0.539	0.358
		0.8	0.851	0.851	0.851	0.841	0.841	0.819	0.838	0.841	0.761	0.575

Table 5.10

Relative Frequency of Rejecting H_0 (LAV fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	ν	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
						n = 8						
PEARSON VII	2.0	0.0	0.072	0.067	0.073	0.075	0.067	0.058	0.124	0.143	0.128	0.077
		0.5	0.277	0.265	0.278	0.286	0.266	0.241	0.388	0.436	0.420	0.287
		0.8	0.616	0.599	0.618	0.631	0.586	0.550	0.720	0.758	0.765	0.581
	1.25	0.0	0.103	0.099	0.103	0.102	0.092	0.084	0.338	0.356	0.173	0.083
		0.5	0.329	0.320	0.330	0.317	0.296	0.273	0.552	0.563	0.406	0.195
		0.8	0.621	0.609	0.622	0.606	0.574	0.553	0.760	0.777	0.672	0.372
						n = 20						
	2.0	0.0	0.074	0.068	0.071	0.057	0.056	0.050	0.195	0.207	0.101	0.072
		0.5	0.623	0.604	0.620	0.563	0.550	0.508	0.655	0.665	0.567	0.468
		0.8	0.980	0.976	0.979	0.961	0.960	0.949	0.937	0.940	0.896	0.802
	1.25	0.0	0.119	0.109	0.116	0.067	0.066	0.053	0.414	0.417	0.162	0.073
		0.5	0.596	0.585	0.593	0.478	0.468	0.435	0.660	0.662	0.422	0.206
		0.8	0.941	0.937	0.940	0.888	0.885	0.874	0.838	0.840	0.688	0.379

Table 5.11

Relative Frequency of Rejecting H_o (OLS fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 8						
BIVARIATE CAUCHY	0.0	0.158	0.152	0.159	0.160	0.147	0.109	0.228	0.247	0.219	0.152
	0.5	0.413	0.387	0.415	0.402	0.369	0.319	0.489	0.516	0.508	0.384
	0.8	0.722	0.691	0.724	0.696	0.649	0.583	0.767	0.792	0.783	0.653
					n = 20						
	0.0	0.206	0.199	0.205	0.185	0.178	0.147	0.322	0.327	0.180	0.113
	0.5	0.658	0.646	0.656	0.588	0.567	0.497	0.654	0.661	0.527	0.401
	0.8	0.926	0.921	0.924	0.904	0.900	0.870	0.879	0.881	0.817	0.682

Table 5.12

Relative Frequency of Rejecting H_0 (LAV fit)
Elliptically Symmetric Model

(nominal α=0.05)

Distribution of (E,E')	λ	T_1	T_2	T_3	K_1^*	K_2^*	K_3^*	R_1	R_2	R_3	R_4
					n = 8						
	0.0	0.091	0.084	0.092	0.086	0.076	0.066	0.211	0.229	0.154	0.087
	0.5	0.312	0.302	0.313	0.308	0.282	0.265	0.462	0.489	0.411	0.253
	0.8	0.634	0.621	0.636	0.636	0.597	0.562	0.733	0.772	0.726	0.499
BIVARIATE CAUCHY					n = 20						
	0.0	0.099	0.086	0.097	0.066	0.065	0.053	0.316	0.321	0.124	0.076
	0.5	0.613	0.503	0.612	0.524	0.513	0.484	0.657	0.666	0.471	0.319
	0.8	0.963	0.958	0.961	0.933	0.927	0.917	0.877	0.879	0.778	0.591

5.4 Conclusions and Recommendations

From tables 5.3-5.6 we can see that for the hypothesis of conditional independence and under the "trivariate reduction" model the performance of Pearson's partial correlation coefficient (R_1 in particular) is very remarkable. For both the OLS fit and the LAV fit and for both small and large samples the test R_1 exhibits an unexpectedly high degree of robustness in terms of both size and power. This is perhaps due to the fact that the "trivariate reduction" model induces a linear structure between E and E', which is the type of structure which occurs in the normal theory models for which Pearson's statistic is designed. For n=20 and for heavy-tailed distributions such as the Cauchy the tests based on T_n have slightly higher powers, but this is perhaps due to their inflated α-levels (see tables 5.5 and 5.6).

For the null hypothesis that $\tau=0$, and for very light-tailed distributions such as the Pearson II (see tables 5.7 and 5.8) the performances of the tests based on R_n are again superior to those of the other tests. For n=20 under the OLS fit, and for both n=8 and n=20 in the case of the LAV fit the tests R_1 and R_2 are conservative (have low α-levels) for <u>very</u> light-tailed distributions (the Pearson II with $\nu=1.0$). In such cases the test R_4 performs the best overall. However, due to the difficulty involved in calculating the statistic R_4 and since in practice one is not usually certain how light tailed the underlying distribution is, a statistic such as R_1 seems to be a better choice.

For medium to heavy-tailed distributions and for testing the null hypothesis that $\tau=0$, tables 5.9-5.12 indicate that tests based on R_n have highly inflated α-levels, low powers or both. The best overall performance in terms of both size and power is that of the test K_3^* which uses the null distribution of K_n^* given in tables 5.1 and 5.2. However, under very heavy-tailed distributions such as the Pearson VII with $\nu=1.25$, and with the OLS fit (see table 5.9) the test K_3^* has highly inflated levels. Since in practice one may have no prior knowledge of the degree of the tail weight of the underlying distribution it is recommended that the LAV estimation be used in testing $\tau=0$.

The summary of our recommendations for testing for partial correlation is as follows.

1) For the hypothesis of conditional independence, and for the hypothesis $\tau=0$ when (E,E') have a very light-tailed distribution, we recommend the use of the usual test (R_1) based on Pearson's partial correlation coefficient.
2) For the hypothesis H_0: $\tau=0$, and for medium to heavy-tailed distributions we recommend the use of the statistic K_n^* obtained from the residuals of a LAV fit and compared to the cut-off values given in table 5.2. For large sample sizes ($n>20$) we suggest comparing K_n^* to the appropriate critical values of the standard normal distribution.

5.5 Related Topics for Future Research

This work is complete only in the sense of fulfilling our initial objective of studying the partial correlation problem under the simple linear setting. However, there are several related problems which need particular attention in future investigations. For example, one may study the partial correlation problem when each of Y and Z are related to a p-variate vector $\underset{\sim}{X}$ by the general linear model or by some other non-linear or functional form. For the simple linear setting one may investigate classes of dependence alternatives other than the "trivariate reduction" model, although our experience shows that this by no means in an easy task as far as theory is concerned.

Another problem of interest is to study the theoretical properties of the statistics K_{RS}^{*} and K_{n}^{*} proposed for testing the null hypothesis that $\tau=0$. For example, one may study the efficiencies of such tests relative to the other tests discussed in this work or investigate their empirical performances under bivariate distributions other than the elliptically symmetric distributions considered here.

APPENDIX
COMPUTER PROGRAM

```
C
C                         PROGRAM ONE
C
C      A PROGRAM TO FIND THE NULL DISTRIBUTION OF THE
C      STATISTIC K* FROM THE L.A.V. FIT OF THE LINEAR
C      MODELS:
C               Y = X + E
C               Z = X + F ,
C      WHERE X HAS THE N(0,1) DISTRIBUTION AND (E,F)
C      HAS THE BIVARIATE PEARSON VII DISTRIBUTICN
       WITH ASSOCIATION PARAMETER RHO=0.0 AND SHAPE
C      PARAMETER NU=2.0:
C
C
C
       INTEGER NR,NS
       REAL R(60)
       REAL X(20),Y(20), Z(20), CC(20)
       REAL E(20),S(20),U(20),V(20)
       REAL W1(20), W2(20), W3(20)
       INTEGER ID(4500)
        INTEGER IND(20), ITER
        REAL FF(20), WT(20), A,B,A1,B1,A2,B2
       DOUBLE PRECISION DSEED1,DSEED2
C
       M=10000
       XM=FLOAT(M)
C
       NR=60
       NS=20
       DSEED1=145645.D0
       DSEED2=123457.D0
C
       DO 9 J=1,4500
9      ID(J)=0
C
       DO 777 K=1,M
C
C
C
       CALL GGNML  (DSEED1,NS,S)
       CALL GGUBS  (DSEED2,NR,R)
       N=20
       XN=FLOAT(N)
C
       NC2=XN*(XN-1.0)/2.0
       FNC2=FLOAT(NC2)
C
C
C
       RHO=0.0
       RHO1=SQRT(1.0-RHO**2)
       EX=-1.0
```

```
      DO 25 I=1,N
      X(I)=S(I)
      I1=2*I-1
      I2=I1+1
      W1(I)=(SQRT((R(I1)**EX)-1.0))
     *     *COS(2.0*3.1416*R(I2))
      W2(I)=(SQRT((R(I1)**EX)-1.0))
     *     *SIN(2.0*3.1416*R(I2))
      W3(I)=RHO*W1(I)+RHO1*W2(I)
      Y(I)=X(I)+W1(I)
      Z(I)=X(I)+W3(I)
25    CONTINUE
C
      CALL DESL1 (Y,X,N,A1,B1,ITER,FF,WT,IND,IFAULT)
      CALL DESL1 (Z,X,N,A2,B2,ITER,FF,WT,IND,IFAULT)
C
      DO 17 I=1,N
         U(I)=Y(I)-A1-B1*X(I)
         V(I)=Z(I)-A2-B2*X(I)
17    CONTINUE
C
C
      CALL TAUHAT (N,XN,U,V,SUMC,SSC)
      TAU=SUMC/(XN*(XN-1.0))
C
      IF (TAU.EQ.1.0) TAU=0.999
      IF (TAU.EQ.-1.0) TAU=-0.999
C
      ZETA1=SSC/(XN*(XN-1.0)*(XN-1.0))
      ZETA2=1.0-TAU*TAU
      ESTVAR=(2.0*(XN-2.0)*ZETA1 + ZETA2)/FNC2
C
      ZTAU=TAU/SQRT(ESTVAR)
C
C
      T=ZTAU+0.0005
      IT=INT(1000*T)
      IF (IT.LT.1000) GO TO 111
      IF (IT.GT.4499) GO TO 222
C
      ID(IT)=ID(IT)+1
      GO TO 777
C
111   ID(999)=ID(999)+1
      GO TO 777
222   ID(4500)=ID(4500)+1
C
C
C
777   CONTINUE
C
      IQ=0
C
      DO 333 J= 999,4500
```

```
          IQ=IQ+ID(J)
          WRITE (6,444) J,IQ
444       FORMAT (' ',2I10)
333    CONTINUE
       STOP
       END
C
C
```

```
C
C                         PROGRAM TWO
C
C      A PROGRAM WHICH CALCULATES THE FREQUENCIES OF
C      REJECTING THE NULL HYPOTHESIS OF THE INDEPEND.
C      BETWEEN THE ERROR TERMS E AND F.  THESE
C      FREQUENCIES ARE CALCULATED FOR THE SEVEN
C      STATISTICS UNDER CONSIDERATION. THIS PROGRAM
C      ALSO CALCULATES THE NUMBER OF TIMES ANY TWO
C      OF THESE STATISTICS REJECT THE NULL
C      SIMULTANEOUSLY:
C
      INTEGER NQ, NR, NS
      REAL R(32), X(8), W1(8), W2(8), W3(8)
      REAL Y(8), Z(8), CC(8),S(8)
      REAL U(8),V(8)
      REAL E1(8), E2(8)
      REAL SS(8),ST(8),TT(8),RST(8),STN(8),PRN(8)
      REAL PTN(8)
      REAL WK(96)
      INTEGER MAT(10,10)
      LOGICAL A,B,C,C1,C2,C3,D,E,F,G
      DOUBLE PRECISION DSEED1,DSEED2,DSEED3
C
C
      M=1000
      XM=FLOAT(M)
C
C
      NR=32
      NQ=1
      NS=8
      DSEED1=143547.D0
      DSEED2=123457.D0
      DSEED3=154677.D0
C
      DO 11 I=1,10
         DO 22 J=1,10
22       MAT(I,J)=0
11    CONTINUE
C
C
      DO 777 K=1,M
C
C
      CALL GGNML (DSEED3,NS,S)
      CALL GGNML (DSEED2,NR,R)
      N=8
      XN=FLOAT(N)
      FNC2=XN*(XN-1.0)/2.0
C
      DELTA=1.7
C
```

```
      DO 25 I=1,N
      X(I)=S(I)
      W1(I)=R(I)
      J=N+I
      W2(I)=R(J)
      JJ=2*N+I
      W3(I)=R(JJ)
      E1(I)=W1(I)+DELTA*W3(I)
      E2(I)=W2(I)+DELTA*W3(I)
      Y(I)=X(I)+E1(I)
      Z(I)=X(I)+E2(I)
25    CONTINUE
C
C
      CALL BETA (N,XN,X,Y,Z,BHAT1,BHAT2)
      DO 27 I=1,N
         U(I)=Y(I)-BHAT1*X(I)
         V(I)=Z(I)-BHAT2*X(I)
27    CONTINUE
C
C
      CALL TAUHAT (N,XN,U,V,SUMC,SSC)
      TAUK=SUMC/2.0
C
      TAU=SUMC/(XN*(XN-1.0))
C
      IF (TAU.EQ.1.0) TAU=0.999
      IF (TAU.EQ.-1.0) TAU=-0.999
      ZETA1=SSC/(XN*(XN-1.0)*(XN-1.0))
      ZETA2=1.0-TAU*TAU
      VARHAT=(2.0*(XN-2.0)*ZETA1 + ZETA2)/FNC2
C
      STARK=TAU/SQRT(VARHAT)
C
C     COMPARE KENDALL'S TAU CALCULATED ON THE RESIDUALS
C     ADJUSTED BY O.L.S. ESTIMATORS TO TABLE A.21 OF
C     HOLLANDER & WOLFE, TO OUR SIMULATED TABLES, AND
C     TO THE Z-TABLES AFTER STANDARDIZATION BY VARIANCE
C     UNDER INDEPENDENCE:
C
C
      IF (TAUK.EQ.14.0) CALL GGUBS (DSEED1,NQ,Q)
      A=  (TAUK.GE.16.0.OR.(TAUK.EQ.14.0.AND.
     *  Q.LE.0.826087))
C
      B= (TAUK.GE.16.0.OR.(TAUK.EQ.14.0.AND.
     *  Q.LE.0.284697))
C
      C= (TAUK.GE.16.0.OR.(TAUK.EQ.14.0.AND.
     *  Q.LE.0.83408))
C
C
C
      IF (A) MAT(1,1)=MAT(1,1)+1
```

```
      IF (B) MAT(2,2)=MAT(2,2)+1
      IF (C) MAT(3,3)=MAT(3,3)+1
      IF (A.AND.B) MAT(1,2)=MAT(1,2)+1
      IF (A.AND.C) MAT(1,3)=MAT(1,3)+1
      IF (B.AND.C) MAT(2,3)=MAT(2,3)+1
C
C
C     COMPARE K* TO THE SIMULATED NULL DISTNS:
C     1) FROM THE ORDINARY CORR. PROBLEM
C     2) FROM OLS RESIDUALS (NORMALITY)
CC    3) FROM OLS RESIDUALS (PEARSON VII):
C
      C1= (STARK.GE.1.78)
      C2= (STARK.GE.1.98)
      C3= (STARK.GE.2.437)
C
      IF (C1) MAT(4,4)=MAT(4,4)+1
      IF (C2) MAT(5,5)=MAT(5,5)+1
      IF (C3) MAT(6,6)=MAT(6,6)+1
      IF (A.AND.C1) MAT(1,4)=MAT(1,4)+1
      IF (A.AND.C2) MAT(1,5)=MAT(1,5)+1
      IF (A.AND.C3) MAT(1,6)=MAT(1,6)+1
      IF (B.AND.C1) MAT(2,4)=MAT(2,4)+1
      IF (B.AND.C2) MAT(2,5)=MAT(2,5)+1
      IF (B.AND.C3) MAT(2,6)=MAT(2,6)+1
      IF (C.AND.C1) MAT(3,4)=MAT(3,4)+1
      IF (C.AND.C2) MAT(3,5)=MAT(3,5)+1
      IF (C.AND.C3) MAT(3,6)=MAT(3,6)+1
      IF (C1.AND.C2) MAT(4,5)=MAT(4,5)+1
      IF (C1.AND.C3) MAT(4,6)=MAT(4,6)+1
      IF (C2.AND.C3) MAT(5,6)=MAT(5,6)+1
C
C
C     COMPARE PEARSON'S R WITH STUDENT-T WITH N-3 DF,
C
      CALL JACK (N,XN,U,V,SYY,SZZ,SYZ,RYZ,TN,VRN,VTN)
C
      IF (RYZ.GE.0.999) RYZ=0.99
C
      RNT=RYZ*SQRT((XN-3.0)/(1.0-RYZ**2))
      D= (RNT.GE.2.015)
C
C
      IF (D) MAT(7,7)=MAT(7,7)+1
      IF (A.AND.D) MAT(1,7)=MAT(1,7)+1
      IF (B.AND.D) MAT(2,7)=MAT(2,7)+1
      IF (C.AND.D) MAT(3,7)=MAT(3,7)+1
      IF (C1.AND.D) MAT(4,7)=MAT(4,7)+1
      IF (C2.AND.D) MAT(5,7)=MAT(5,7)+1
      IF (C3.AND.D) MAT(6,7)=MAT(6,7)+1
C
C     COMPARE THE TRANSFORMED FISHER'S Z
C     STANDARDIZED BY ITS VARIANCE 1/N-3 TO Z_0.05:
C
```

```
      ZTN=SQRT((XN-3.0))*TN
      E= (ZTN.GE.1.645)
C
      IF (E) MAT(8,8)=MAT(8,8)+1
      IF (A.AND.E) MAT(1,8)=MAT(1,8)+1
      IF (B.AND.E) MAT(2,8)=MAT(2,8)+1
      IF (C.AND.E) MAT(3,8)=MAT(3,8)+1
      IF (C1.AND.E) MAT(4,8)=MAT(4,8)+1
      IF (C2.AND.E) MAT(5,8)=MAT(5,8)+1
      IF (C3.AND.E) MAT(6,8)=MAT(6,8)+1
      IF (D.AND.E) MAT(7,8)=MAT(7,8)+1
C
C
C     USE THE JACKKNIFE ESTIMATORS OF THE VARIANCES OF
C     OF PEARSON'S R, AND FISHER'S Z, AND COMPARE TO
C     Z_0.05:
C
C
      SDRN=SQRT(VRN/XN)
      SDTN=SQRT(VTN/XN)
      RNJACK=RYZ/SDRN
      TNJACK=TN/SDTN
C
      F= (RNJACK.GE.1.645)
      G= (TNJACK.GE.1.645)
C
      IF (F) MAT(9,9)=MAT(9,9)+1
      IF (G) MAT(10,10)=MAT(10,10)+1
      IF (A.AND.F) MAT(1,9)=MAT(1,9)+1
      IF (B.AND.F) MAT(2,9)=MAT(2,9)+1
      IF (C.AND.F) MAT(3,9)=MAT(3,9)+1
      IF (C1.AND.F) MAT(4,9)=MAT(4,9)+1
      IF (C2.AND.F) MAT(5,9)=MAT(5,9)+1
      IF (C3.AND.F) MAT(6,9)=MAT(6,9)+1
      IF (D.AND.F) MAT(7,9)=MAT(7,9)+1
      IF (E.AND.F) MAT(8,9)=MAT(8,9)+1
      IF (A.AND.G) MAT(1,10)=MAT(1,10)+1
      IF (B.AND.G) MAT(2,10)=MAT(2,10)+1
      IF (C.AND.G) MAT(3,10)=MAT(3,10)+1
      IF (C1.AND.G) MAT(4,10)=MAT(4,10)+1
      IF (C2.AND.G) MAT(5,10)=MAT(5,10)+1
      IF (C3.AND.G) MAT(6,10)=MAT(6,10)+1
      IF (D.AND.G) MAT(7,10)=MAT(7,10)+1
      IF (E.AND.G) MAT(8,10)=MAT(8,10)+1
      IF (F.AND.G) MAT(9,10)=MAT(9,10)+1
C
C
777   CONTINUE
C
C
      DO 888 I=1,10
         WRITE (6,123) (MAT(I,J), J=1,10)
123      FORMAT ('-',10I8)
888   CONTINUE
```

```
C
      STOP
      END
```

```
C
C                        SUBROUTINE ONE
C
C     THIS SUBROUTINE CALCULATES THE LEAST ABSOLUTE
C     VALUE (LAV) PARAMETER ESTIMATES OF THE SIMPLE
C     LINEAR MODEL (JOSVANGER AND SPOSITO, 1983):
C
C
      SUBROUTINE DESL1(Y,X,N,A,B,ITER,FF,WT,IND,IFAULT)
C
      INTEGER IND(N), ITER
      REAL X(20), Y(20), FF(20), WT(20), A, B,A1,A2,B1
      REAL B2
      DATA TOL/1.0E-6/
      XN=FLOAT(N)
C
C     FIND ESTIMATES OF A AND B
C
      IFAULT=0
      SY=0.0
      A1=X(1)
      DO 10 I=1,N
         IF (X(I).NE.A1) IFAULT=1
         SY=SY+Y(I)
10    CONTINUE
      B=0.
      A=SY/XN
      IF (IFAULT.EQ.0) RETURN
      ITER=0
      IRESA=1
      IRESB=0
      DEV=ABS(Y(1)-A)
      DO 20 J=2,N
         IF (ABS(Y(J)-A).GE.DEV) GO TO 20
         DEV=ABS(Y(J)-A)
         IRESA=J
20    CONTINUE
C
C     RECORD X AND Y CORRESP. TO THE MIN. ABS. DEVIATION
C
30    J=IRESA
      XJ=X(J)
      YJ=Y(J)
      I=1
      M=1
      K=N
      TWL=0.
      TWU=0.
C
C     SEPERATE SLOPE VALUES > B FROM THOSE < B
C
40    XIJ=(X(I)-X(J))
      IF (XIJ.EQ.0.) GO TO 60
```

```
      FI=(Y(I)-YJ)/XIJ
      IF (XIJ.LT.0.) XIJ=-XIJ
      IF (FI.GT.B) GO TO 50
      FF(M)=FI
      WT(M)=XIJ
      IND(M)=I
      TWL=TWL+XIJ
      M=M+1
      GO TO 60
50    FF(K)=FI
      WT(K)=XIJ
      IND(K)=I
      TWU=TWU+XIJ
      K=K-1
60    IF (I.EQ.N) GO TO 70
      I=I+1
      GO TO 40
C
C     SET THE NEW B VALUE = WEIGHTED MEDIAN SLOPE
C
70    ASUM=(TWL+TWU)/2.
      IF (TWL.GE.TWU) GO TO 130
      M=M-1
80    K=K+1
      M=M+1
      I=K
90    FNEW=FF(I)
      INEW=I
100   IF (I.EQ.N) GO TO 110
      I=I+1
      IF (FF(I).LT.FNEW) GO TO 90
      GO TO 100
110   TWL=TWL+WT(INEW)
      IF (TWL.GE.ASUM) GO TO 120
      FF(INEW)=FF(K)
      WT(INEW)=WT(K)
      IND(INEW)=IND(K)
      GO TO 80
120   BNEW=FNEW
      JNEW=IND(INEW)
      GO TO 180
130   M=M-1
      I=M
140   FNEW=FF(I)
      INEW=I
150   IF (I.EQ.1) GO TO 160
      I=I-1
      IF (FF(I).GT.FNEW) GO TO 140
      GO TO 150
160   TWU=TWU+WT(INEW)
      IF (TWU.GT.ASUM) GO TO 170
      FF(INEW)=FF(M)
      WT(INEW)=WT(M)
      IND(INEW)=IND(M)
```

```
      GO TC 130
 170  BNEW=PNEW
      JNEW=IND(INEW)
C
C     FIND NEW INTERCEPT VALUE
C     CHANGE ITERATICN COUNT
C
180   ITER=ITER+1
      A=YJ-BNEW*XJ
C
C     TEST ONE FOR SOLUTION:
C     COMPARE DIFFERENCE IN B VALUES TO TOLERANCE
C     LEVEL
C
      IF (ABS(B-BNEW).LE.TOL) GO TO 190
      B=BNEW
C
C     TEST TWO FOR SOLUTION:
C     CHECK FOR REPITITION IN INDEX PATTERN
C
      IF (IRESB.EQ.JNEW) GO TO 190
      IRESB=IRESA
      IRESA=JNEW
      GO TC 30
190   RETURN
      END
```

```
C
C                          SUBROUTINE TWO
C
C     THIS SUBROUTINE CALCULATES KENDALL'S TAU. THE
C     NUMBERS OBTAINED THROUGH THIS SUBROUTINE ARE
C     USED IN THE MAIN PROGRAM TO FIND K*:
C
      SUBROUTINE TAUMAT (N,XN,Y,Z,SUMC,SSC)
      DIMENSION Y(8), Z(8),CC(8)
      SUMC=0.0
      SSC=0.0
C
      DO 1 I=1,N
      COUNT=0.0
         DO 2 J=1,N
           YIJ=Y(I)-Y(J)
           ZIJ=Z(I)-Z(J)
           IF ((YIJ*ZIJ).GT.0.0) COUNT=COUNT+1.0
           IF ((YIJ*ZIJ).LT.0.0) COUNT=COUNT-1.0
2        CONTINUE
         CC(I)=COUNT
         SUMC=SUMC+CC(I)
C
1     CONTINUE
C
      CBAR=SUMC/XN
C
      DO 3 I=1,N
3     SSC=SSC+(CC(I)-CBAR)*(CC(I)-CBAR)
C
      RETURN
      END
```

```
C
C                             SUBROUTINE THREE
C
C
      SUBROUTINE JACK (N,XN,Y,Z,SYY,SZZ,SYZ,RYZ,
     *  TN,VRN,VTN)
      DIMENSION Y(20), Z(20)
      DOUBLE PRECISION DXN,DY(20),DZ(20),DSYY,DSZZ
      DOUBLE PRECISION DVRN,DVTN,SS(20),ST(20)
      DOUBLE PRECISION STN(20),PRN(20),PTN(20),DTN
      DOUBLE PRECISION SUM3,SUM4,SUMPRN,SUMPTN
      DOUBLE PRECISION BARY,BARZ
      DOUBLE PRECISION DSYZ,DRYZ,TT(20),RST(20)
      DOUBLE PRECISION SUM1,SUM2,SAVEY,SAVEZ,S2, T2
      DOUBLE PRECISION SUMY,SUMZ,YBAR,ZBAR
C
      SUMY=0.D0
      SUMZ=0.D0
      DSYY=0.D0
      DSZZ=0.D0
      DSYZ=0.D0
      SUM3=0.D0
      SUM4=0.D0
      SUMPRN=0.D0
      SUMPTN=0.D0
      DXN=XN
      DO 5 I=1,N
       DY(I)=Y(I)
       DZ(I)=Z(I)
        SUMY=SUMY+DY(I)
        SUMZ=SUMZ+DZ(I)
5     CONTINUE
      YBAR=SUMY/DXN
      ZBAR=SUMZ/DXN
         DO 6 I=1,N
         DSYY=DSYY+(DY(I)-YBAR)*(DY(I)-YBAR)
         DSZZ=DSZZ+(DZ(I)-ZBAR)*(DZ(I)-ZBAR)
         DSYZ=DSYZ+(DY(I)-YBAR)*(DZ(I)-ZBAR)
6        CONTINUE
C
      DRYZ=DSYZ/DSQRT(DSYY*DSZZ)
      DTN=0.5D0*(DLOG(1.D0+DRYZ)-DLOG(1.D0-DRYZ))
C
      RYZ=DRYZ
C
      DO 10 I=1,N
      SAVEY=DY(I)
      S2=SAVEY
      DY(I)=0.D0
      SAVEZ=DZ(I)
      T2=SAVEZ
      DZ(I)=0.D0
      SUM1=0.D0
```

```
      SUM2=0.D0
C
      DO 11 J=1,N
         SUM1=SUM1+DY(J)
         SUM2=SUM2+DZ(J)
11    CONTINUE
C
      DY(I)=SAVEY
      DZ(I)=SAVEZ
C
      BARY=SUM1/(DXN-1.D0)
      BARZ=SUM2/(DXN-1.D0)
      SS(I)=DSYY-((DXN-1.D0)/DXN)*(BARY-S2)*(BARY-S2)
      TT(I)=DSZZ-((DXN-1.D0)/DXN)*(BARZ-T2)*(BARZ-T2)
      ST(I)=DSYZ-((DXN-1.D0)/DXN)*(BARY-S2)*(BARZ-T2)
      RST(I)=ST(I)/DSQRT(SS(I)*TT(I))
      STN(I)=0.5D0*(DLOG(1.D0+RST(I))-DLOG(1.D0-
     *   RST(I)))
      PRN(I)=DXN*DRYZ-(DXN-1.D0)*RST(I)
      PTN(I)=DXN*DTN-(DXN-1.D0)*STN(I)
      SUMPRN=SUMPRN+PRN(I)
      SUMPTN=SUMPTN+PTN(I)
C
10    CONTINUE
C
      PRNBAR=SUMPRN/DXN
      PTNBAR=SUMPTN/DXN
      DO 12 J=1,N
      SUM3=SUM3+(PRN(J)-PRNBAR)*(PRN(J)-PRNBAR)
      SUM4=SUM4+(PTN(J)-PTNBAR)*(PTN(J)-PTNBAR)
12    CONTINUE
      DVRN=SUM3/(DXN-1.D0)
      DVTN=SUM4/(DXN-1.D0)
C
      TN=DTN
      VRN=DVRN
      VTN=DVTN
      RETURN
      END
C
```

BIBLIOGRAPHY

Agresti, A. and Agresti, B.F. (1979). Statistical Methods for the Social Sciences. San Francisco: Dellen Publishing Company.

Anderson, T.W. (1958). An Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons.

Bassett, G., Jr., and Koenker, R. (1978). Asymptotic Theory of Least Absolute Error Regression. J. Amer. Stat. Assoc., 73, 618-622.

Bhuchongkul, S. (1964). A Class of Nonparametric Tests for Independence in Bivariate Populations. Ann. Math. Statist., 35, 138-149.

Cramér, H. (1946). Mathematical Methods of Statistics. Princeton: Princeton University Press.

Fligner, M.A. and Rust, S.W. (1983). On the Independence Problem and Kendall's Tau. Comm. in Stat., A12, 1597-1607.

Gokhale, D.V. (1966). Some Problems in Independence and Dependence. Ph.D. Dissertation, Department of Statistics, University of California at Berkeley, Berkeley, California.

Hájek, J. and Sidák, Z. (1967). The Theory of Rank Tests. New York: Academic Press.

Hinkley, D.V. (1977). Jackknife Confidence Limits Using Student t Approximations. Biometrika, 64, 21-28.

__________ (1978). Improving the Jackknife with Special Reference to Correlation Estimation. Biometrika, 65, 13-21.

Hollander, M. and Wolfe, A.W. (1973). Nonparametric Statistical Methods. New York: John Wiley and Sons.

Johnson, M.E. and Ramberg, J.S. (1977). Elliptically Symmetric Distributions: Characterization and Random Variate Generation. Stat. Comp. Sect. Proc. of ASA, 262-265.

Josvanger, L.A. and Sposito, V.A. (1983). L_1 Norm Estimates for the Simple Regression Problem. Comm. in Stat., B12, 215-221.

Kendall, M.G. (1962). Rank Correlation Methods. New York: Hafner.

Kepner, J.L. (1979). Tests Using the Null Hypothesis of Bivariate Symmetry. Ph.D. Dissertation in Statistics, University of Iowa, Iowa City.

Konijn, H.S. (1956). On the Power of Certain Tests for Independence in Bivariate Populations. Ann. Math. Statist., 27, 300-323.

Korn, E.L. (1984). The Ranges of Limiting Value of Some Partial Correlations Under Conditional Independence. The American Statistician, 38, 61-62.

Magsoodloo, S. (1975). Estimates of the Quantiles of Kendall's Partial Rank Correlation Coefficient. J. Stat. Comput. Simul., 4, 155-164.

__________ and Pallos, L.L. (1981). Asymptotic Behavior of Kendall's Partial Rank Correlation Coefficient and Additional Quantile Estimates. J. Stat. Comput. Simul., 13, 41-48.

Miller, R.G. (1974). The Jackknife--A Review. Biometrika, 61, 1-15.

Noether, G.E. (1955). On a Theorem of Pitman. Ann. Math. Stat., 26, 64-68.

Ott, L., Larson, R.F. and Mendenhall, W. (1983). Statistics: A Tool for the Social Sciences. Boston: Duxbury Press.

Quade, D. (1974). Nonparametric Partial Correlations, in Measurements in the Social Sciences, ed. H.M. Blalock, Jr. Chicago: Aldine Publishing.

Randles, R.H. (1982). On the Asymptotic Normality of Statistics with Estimated Parameters. Annals of Stat., 10, 462-474.

__________ (1984). On Tests Applied to Residuals. J. Amer. Stat. Assoc., 79, 349-354.

__________ and Wolfe, D.A. (1979). Introduction to the Theory of Nonparametric Statistics. New York: John Wiley and Sons.

Sen, P.K. (1960). On Some Convergence Properties of U-Statistics. Calcutta Statistical Association Bulletin, 10, 1-18.

__________ (1968). Estimates of the Regression Coefficient Based on Kendall's Tau. J. Amer. Statist. Ass., 63, 1379-1389.

Serfling, R.J. (1980). Approximation Theorems of Mathematical Statistics. New York: John Wiley and Sons.

Shirahata, S. (1977). Tests of Partial Correlation in a Linear Model. Biometrika, 64, 162-164.

Sukhatme, B.V. (1958). Testing the Hypotheses that Two Populations Differ Only in Location. Ann. Math. Statist., 29, 60-78.

BIOGRAPHICAL SKETCH

Basil Samara was born to Lebanese parents in Tehran, Iran, on December 17, 1948. He moved to Lebanon in 1958 where he attended Marj'Oyoun National School in Marj'Oyoun, Lebanon, until 1967. Upon receiving the Lebanese Baccalaureate II from the International College of Beirut in 1968, he attended Kalamazoo College, Michigan, where he received his Bachelor of Arts degree in mathematics in 1971. After graduation, he returned to Beirut, Lebanon, where he taught in several high schools and colleges until 1976. In 1977 he entered Miami University, Ohio, where he received his Master of Statistics degree in 1979, after which he joined the faculty of the Business School of Miami University as an instructor of statistics. In 1981, he entered the graduate program in statistics at the University of Florida, and he received the degree of Doctor of Philosophy in August, 1985. He is a member of Phi Beta Kappa and the American Statistical Association.

His professional career has included teaching mathematics and physics at International College, St. Mary's College and Rawdah High School in Beirut, and teaching mathematics and statistics as a graduate teaching assistant at Miami University and the University of Florida.

He has been the recipient of the Lebanese Government's Brevet, Bacc. I and Bacc. II awards, scholarships and graduate assistantships throughout his academic career. He has also received the Graduate Student Teaching Award at the University of Florida for the year 1985.

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Ronald H. Randles

Ronald Randles, Chairman
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

William C. Parr

William Parr
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Malay Ghosh

Malay Ghosh
Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Dennis D. Wackerly

Dennis Wackerly
Associate Professor of Statistics

I certify that I have read this study and that in my opinion it conforms to acceptable standards of scholarly presentation and is fully adequate, in scope and quality, as a dissertation for the degree of Doctor of Philosophy.

Robert Long.

Robert Long
Assistant Professor of Mathematics

This dissertation was submitted to the Graduate Faculty of the Department of Statistics in the College of Liberal Arts and Sciences and to the Graduate School and was accepted as partial fulfillment of the requirements for the degree of Doctor of Philosophy.

August, 1985

Madelyn Lockhart

Dean, Graduate School